Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters

Andrew W. Delton[a,b,1,2], Max M. Krasnow[a,b,1,2], Leda Cosmides[a,b], and John Tooby[a,c]

[a]Center for Evolutionary Psychology and Departments of [b]Psychology and [c]Anthropology, University of California, Santa Barbara, CA 93106

Are humans too generous? The discovery that subjects choose to incur costs to allocate benefits to others in anonymous, one-shot economic games has posed an unsolved challenge to models of economic and evolutionary rationality. Using agent-based simulations, we show that such generosity is the necessary byproduct of selection on decision systems for regulating dyadic reciprocity under conditions of uncertainty. In deciding whether to engage in dyadic reciprocity, these systems must balance (*i*) the costs of mistaking a one-shot interaction for a repeated interaction (hence, risking a single chance of being exploited) with (*ii*) the far greater costs of mistaking a repeated interaction for a one-shot interaction (thereby precluding benefits from multiple future cooperative interactions). This asymmetry builds organisms naturally selected to cooperate even when exposed to cues that they are in one-shot interactions.

altruism | cooperation | ecological rationality | social evolution | evolutionary psychology

Human behavior in all known cultures is densely interpenetrated by networks of reciprocity or exchange (to use the terms of biologists and economists, respectively). Fueled by these observations, biologists and game theorists developed models that outlined how the fitness benefits to be reaped from gains in trade can, under the right envelope of conditions, drive the evolution of decision-making adaptations for successfully engaging in direct reciprocity (1–5). Indeed, a broad array of experimental and neuroscientific evidence has accumulated over the last two decades supporting the hypothesis that our species' decision-making architecture includes both cognitive and motivational specializations whose design features are specifically tailored to enable gains through direct reciprocity (e.g., detection of defectors and punitive sentiment toward defectors) (6–16).

The most important condition necessary for the evolution of direct reciprocity is that interactions between pairs of agents be sufficiently repeated (2). For reciprocity to operate, after one agent delivers a benefit, the partner must forgo the immediate gain offered by cheating—that is, of not incurring the cost involved in returning a comparable benefit. In general, selection can only favor forgoing this gain and incurring the cost of reciprocating when the net value to the partner of the future series of exchange interactions (enabled by reciprocation) exceeds the benefit of immediate defection (which would terminate that future series). If there were no future exchanges—if an interaction was one-shot—then the equilibrium strategy would be *always defect*. However, both direct observations and the demographic conditions that characterize hunter–gatherer life indicate that large numbers of repeat encounters, often extending over decades, was a stable feature of the social ecology of ancestral humans (17).

Despite this close fit between theory and data for direct reciprocity, problems emerged in closely related issues. In particular, when experimentalists began using laboratory economic games to test theories of preferences and cooperation, they uncovered some serious discrepancies between observed experimental behavior and the predictions of traditional economic models of rationality and self-interest (18–21). Some of these results have

proven equally challenging to biologists, because they seem to violate the expectations of widely accepted models of fitness maximization that predict selfishness in the absence of either (*i*) genetic relatedness, (*ii*) conditions favoring reciprocity, or (*iii*) reputation enhancement (22).

The most glaring anomaly stems from the fact that, according to both evolutionary and economic theories of cooperation, whether an interaction is repeated or one-shot should make a crucial difference in how agents act. The chance of repeated interactions offers the only possibility of repayment for forgoing immediate selfish gains when the interactants are not relatives and the situation precludes reputation enhancement in the eyes of third parties. Specifically, in an anonymous one-shot interaction, where (by definition) there will be no future interactions or reputational consequences, it seems both well-established and intuitive that rational or fitness-maximizing agents ought to choose the higher payoff of behaving selfishly (defecting) over cooperating. In one-shot games, cooperative or other altruistic choices were theoretically expected to vanish (2, 23).

Empirically, however, individuals placed in anonymous, one-shot experimental games seem far more altruistic than biologists, economists, and game theorists predicted (18–22). To explain these anomalies, a proliferating series of alternative economic, evolutionary, cultural, and psychological explanations has been advanced, many proposing major revisions to the foundations of economic theory and to standard views of how social evolution typically works (24–30). These explanations have ranged from reconceptualizations of economic rationality and proposals of generalized other-regarding preferences to accounts of altruism produced variously through genetic group selection, cultural group selection, or gene–culture coevolution.

Whatever the merit of these hypotheses (31–36), however, they all start from the widely accepted assumption that the highest paying (or most rational) decision procedure is: *If you are in a one-shot interaction, always defect*. However, a number of questions arise when one dissects this assumption as part of an attempt to turn it into decision-making procedures that an evolved agent operating in the real world could actually carry out: How much evidence should ideally be required before categorizing an interaction as one-shot? Given uncertainty about this categorization, is *Always defect when the interaction has been categorized as one-shot* truly the payoff-maximizing strategy? We argue that once these questions are addressed, it becomes clear that the strategy *If you are in a one-shot interaction, always defect* is either defective or

PSYCHOLOGICAL AND COGNITIVE SCIENCES

inapplicable to evolved cooperative architectures, such as those found in the human mind. Using agent-based simulations, we show that a propensity to make contributions in one-shot games (even those without reputational consequences) evolves as a consequence of including in the architecture an overlooked computational step necessary for guiding cooperative decisions: the discrimination of one-shot from repeated interactions.

One-shot discrimination, like cheater detection (37), enables cooperative effort to be directed away from unproductive interactions. However, to behave differently in one-shot vs. repeated interactions requires the capacity to distinguish them, a judgment that must be made under uncertainty. These simulations explore the impact on cooperative decision-making architectures (or equivalently, on cooperative strategies) of evolving in conditions of uncertainty—conditions in which the discrimination of one-shot from repeated interactions can be better or worse, but never perfect and error-free.*

Imperfect discrimination is the biologically realistic case because real computational systems, such as human minds, cannot know with certainty whether an interaction is one-shot at the time the decision about cooperating must be made. Indeed, given the stochastic nature of the world, it might be correct to say that, at the time of the interaction, the interaction is not determinately either one-shot or repeated. Instead, an interaction only becomes one-shot retroactively at events that uniquely preclude additional interactions, such as the death of one of the parties. Of course, certain situations or persons may exhibit cues that lead a decision-maker to judge that an interaction is highly likely to be one-shot, but that probability can never reach certainty: While both parties live, there is always a nonzero probability of a repeated encounter, whether through error, intent, or serendipity. This logic applies with special force to the small-scale world of our ancestors, where travel was on foot and population sizes were small. In such a world, a first encounter with someone suggests a nonzero probability of encountering them again.

To judge whether an interaction is likely to be repeated or one-shot, decision-making designs engineered by evolution must use probabilistic information available in the social ecology—cues that differentially predict the two types of interactions. These cues can be present in the situation (e.g., you are traveling far from home), in the person encountered (e.g., the interactant speaks with your accent), or in both (e.g., the interactant marries into your band). In the real world, however, no cue will be perfectly predictive, and the presence of several cues may entail conflicting implications (even in laboratory experiments, verbal assurances from an experimenter may compete with situational cues suggesting that other interactants are members of one's community). Consequently, each choice is a bet, and even an ideal Bayesian observer will make errors. Therefore, agents face a standard Neyman–Pearsonian decision problem (38) with two types of possible errors: false positives and misses. A false-positive error occurs when the agent decides that (or acts as if) an interaction will be repeated, but it turns out to be one-shot. A miss occurs when the agent decides that (or acts as if) the interaction will be one-shot, but it turns out to be repeated.

Under uncertainty, decision architectures (or strategies) cannot simultaneously eliminate both types of errors; making fewer errors of one type must be paid for with more errors of the other type. If the two errors inflict costs of different magnitudes, selection will favor a betting strategy that buys a reduction in the more expensive error type with an increase in the cheaper type (13, 39, 40). We show here that the costs to the architecture of the two types of errors will be different under nearly all conditions. Defecting has the modest upside of gaining from a single instance of benefit-withholding (if it is a one-shot interaction) and a downside equal to the summed value of the future series of benefit–benefit interactions precluded by defection (if the interaction turns out to be repeated, but the partner withdraws her cooperative effort in response to the initial defection). In social ecologies where repeat interactions are numerous enough to favor the evolution of reciprocation, the value of such a forfeited benefit series will generally be large compared with expending a single retrospectively unnecessary altruistic act. Overall, these asymmetric costs evolutionarily skew decision-making thresholds in favor of cooperation; as a result, fitter strategies will cooperate "irrationally" even when given strong evidence that they are in one-shot interactions.

Model

To investigate the effect of selection on cooperative decision-making in a world where organisms can only imperfectly discriminate one-shot from repeated interactions, we conducted simulations[†] of agents interacting in two-player prisoner's dilemmas (PDs). (An analytic exploration is in *SI Text*.) PDs, the most common model for cooperative decision-making, can have one or more rounds of interaction. Within a round, defection always pays more than cooperation, regardless of the other player's actions. However, the relative costs and benefits are arranged such that mutual cooperation is better than mutual defection (Fig. 1). Agents face two types of PDs: one-shot PDs consisting of a single round, and indefinitely repeated PDs consisting of a first round and a stochastic number of subsequent rounds. Agents are not given perfect knowledge of which type of PD they are in, but only a probabilistic cue.

Agents can commit two types of errors. One error is to cooperate in a one-shot interaction, paying the cost of cooperating without increasing the benefits received (Fig. 1 *Right*, upper right square). A second error is to defect in a repeated interaction, missing the associated opportunity for long-term, mutually beneficial exchange. This error is typically far more costly. As shown in Fig. 1 *Left*, upper left square, when (*i*) the net within-round benefits of cooperation and (*ii*) the number of rounds become large enough, this payoff will be much larger than any other payoff.

In each simulation, all agents have identical rules for deciding when to cooperate or defect; the only component of their psychology that can vary is the value of a regulatory variable that is consulted by these rules. Before their interactions, they observe a cue value that probabilistically—but imperfectly—reflects the type of PD that they face. The decision rule then operates on the cue with procedures that are calibrated by an evolvable regulatory variable to produce a decision of whether to cooperate or defect. Described more fully below (*SI Text*), these variables determine how likely agents are to cooperate or defect based on the cues that they perceive. We allow the magnitudes of these decision-regulating variables to evolve by natural selection based on their fitness consequences. To explore how broad or narrow the conditions are that favor cooperation in one-shot interactions, across simulations we parametrically varied (*i*) the relative frequency of one-shot vs. repeated interactions, (*ii*) the average length of indefinitely repeated interactions, and (*iii*) the within-round benefits of cooperation.

Our goal is to determine whether agents evolve to cooperate as opposed to defect in one-shot interactions; it is not to determine exactly which of the many cooperative strategies dis-

---

*Of course, for an interaction to qualify as repeated in a way that enables cooperation to evolve, the actor must be able to identify the interactant as being the same individual across repeated interactions. In empirical work, experimental anonymity plays two conceptually distinct roles. It precludes third parties from responding to player behavior, and it prevents partners from recognizing each other across interactions. The models that we present here preclude any third-party responses but presume that interactants can identify each other across repeated interactions.

---

†The simulation program was written in Java by M.M.K. and checked for errors by A.W.D. Source code is available on request to M.M.K.

Fig. 1. Under realistic parameter values, the largest possible benefits occur when both players cooperate in a repeated interaction. This figure shows payoffs in a two-player PD as a function of players' strategies and whether the interaction is one-shot or repeated. Payoffs assume that cooperation entails playing TIT-for-TAT or GRIM, and defection entails defecting forever (*SI Text*). Avg. Rounds, average number of rounds of the PD in repeated interactions; b, within-round benefit delivered to partner by cooperation; c, cost of delivering within-round benefit.

cussed in the literature would ultimately prevail in our simulated social ecology. Thus, as a representative strategy, our main simulations use the well-known, conditionally cooperative TIT-for-TAT strategy to represent an agent's choice to cooperate. TIT-for-TAT cooperates on the first round of an interaction and thereafter copies its partner's behavior from the previous round (2). (Using a contingently cooperative strategy is important because noncontingent, pure cooperation cannot evolve, even if all interactions are indefinitely repeated.) When an agent chooses to cooperate, it does so by playing TIT-for-TAT. When it chooses to defect instead, it defects on all rounds of an interaction, regardless of its partner's behavior. We also checked whether our main results are robust against the possibility of behavioral errors (accidental defections) by using a contingently cooperative strategy that is much less forgiving than TIT-for-TAT, permanently shutting down cooperation after even a single act of defection (see below).

In the real world, agents can perceive multiple cues that probabilistically predict whether their interaction with a partner will be one-shot or repeated. Some of these cues involve personal characteristics of the partner, whereas other cues involve characteristics of the situation. Although the perception and use of such cues is surely an independent target of selection, such complexity is orthogonal to the scope of this investigation. Rather, we simply model this informational ecology by assuming that all cues, situational and personal, could be collapsed into a single number—a cue summary—that is associated with a partner. Cue summaries are modeled as two normally distributed random variables, one for each type of PD (Fig. 2). Larger observed cue summaries imply a greater chance that the interaction is repeated. The fourth and final exogenous parameter varied in
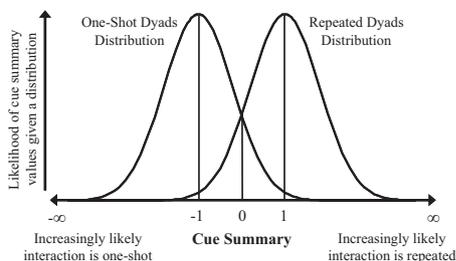

Fig. 2. The distributions of cue summary values showing that repeated interactions are associated with larger-valued cue summaries. Agents are paired and randomly assigned to be part of one-shot or repeated interactions irrespective of genotype. Moreover, given the interaction that they face, agents are randomly assigned cue summaries from the appropriate cue summary distribution, irrespective of their genotype.

our simulations is the distance between the cue summary distributions. A greater distance makes it easier to correctly discriminate one-shot from repeated interactions.

Within a simulation, each generation of 500 agents was sorted into dyads, and dyads were assigned to be one-shot or repeated, both randomly with respect to genotype. Each member of a dyad independently and randomly with respect to genotype drew a cue summary from the one-shot distribution (if they had been assigned to a one-shot dyad) or from the repeated distribution (if they had been assigned to a repeated dyad) (Fig. 2). Thus, agents' strategies (i.e., the evolvable variables embedded in their decision rules) are completely uncorrelated with which type of PD they face and the value of their cue summary. Of course, cue summaries and types of PDs are necessarily—although imperfectly—correlated. Within each generation, agents behaved in accordance with their decision rules, accrued fitness based on their PD's outcome, reproduced asexually in proportion to their fitness, and then died. This cycle occurred through 10,000 generations for each simulation run. Each run had fixed parameter values, and four replicate runs were conducted for each fixed parameter set.

As a clarifying idealization, one can partition a behavior-regulating architecture (or strategy) into components concerned with representing states of the world (cognitive components) and components that transform such representations into decisions about what actions to take (motivational components). The strategy most often articulated in standard models—*Defect when you are in a one-shot interaction*—can be mapped into these components as follows:

i) Cognitive component: Compute your beliefs about whether you are in a one-shot or repeated interaction as accurately as possible.

ii) Motivational component: Given the state of your beliefs, be motivated to act in a manner consistent with your beliefs—that is, act to give yourself the highest payoff, assuming your belief is true. This logic reduces to: *If you believe that you are in a one-shot interaction, defect; If you believe that you are in repeated interaction, cooperate.*

Each of these decision-making components embodies canonically rational decision-making methods, and jointly, they prescribe what is typically believed to be ideal strategic behavior in PDs. Here, we show that neither of these forms of rationality produces the fittest or best-performing strategic choice when there is uncertainty about whether an interaction is one-shot or repeated. That is, when key aspects of these decision-making rules are subjected to mutation and selection, better performing alternatives evolve.

In a first set of simulations, the motivational component is allowed to evolve, whereas the cognitive component is fixed. Here, the agents' cognitive component specifies that they use Bayesian updating to form ideally rational beliefs to discriminate the kind of interaction that they face. The motivational component, however, references an evolvable regulatory variable: *Cooperation Probability$_{One-Shot}$*. This variable specifies the probability of cooperation given a belief that the interaction is one-shot. Instead of being fixed at what is typically assumed to be the optimal value—0% cooperation (i.e., 100% defection)—in these simulations, the probability of cooperation given a one-shot belief is left free to evolve.

In a second set of simulations, the motivational component is fixed such that organisms never cooperate given a belief the interaction is one-shot and always cooperate given a belief the interaction is repeated. Now, the cognitive component references an evolvable regulatory variable, a decision threshold; this variable sets the weight of evidence required for the agent to conclude that it is in a one-shot interaction.

In the human case, of course, both cognitive and motivational elements might have been simultaneously shaped by these selection pressures, leaving an infinite set of candidate cooperative architectures to consider. The two architectures explored here, however, bracket the spectrum of possible architectures, allowing us to separately test the performance of each component of standard rationality (more discussion is in *SI Text*).

Results

Simulation Set 1. Given accurate belief formation, how will selection shape the motivation to cooperate? Starting with a common assumption in economics and psychology, in this set of simulations, agents' decision rules use Bayesian updating to form optimally rational beliefs (23, 41). Using Bayesian updating, agents compute the (posterior) probabilities that their interaction is one-shot or repeated by integrating their partner's cue summary with the relative frequencies of one-shot and repeated PDs in their environment. For simplicity, we assume that our agents have innate and perfect knowledge of these base rates (this works against our hypothesis by giving agents the most accurate beliefs possible). Given these updated probabilities, if it is more likely that the interaction is one-shot, then the agent believes that it is one-shot; otherwise, the agent believes that it is repeated.

Unlike standard models of cooperation, however, there is no direct mapping between belief and action in this simulation. Agents with one-shot beliefs do not necessarily defect; agents with repeated beliefs do not necessarily cooperate. Instead, agents access one of two inherited motivational variables: *Cooperation Probability$_{One-Shot}$* and *Cooperation Probability$_{Repeated}$*. Because these variables are subject to mutation and selection, and hence can evolve, it is possible for selection to move one or both to any value between zero and one.

In contrast to previous accounts, our analysis predicts that *Cooperation Probability$_{One-Shot}$* will evolve to be larger than zero—generating cooperation even when agents have one-shot beliefs—because the costs of missing repeated cycles of mutually beneficial cooperation outweigh the costs of mistaken one-shot cooperation. We test this prediction by setting the regulatory variables of the first generation such that beliefs and actions are perfectly consistent—optimal on the standard view. First-generation agents always cooperate when they believe the interaction is repeated (*Cooperation Probability$_{Repeated}$* ≈ 1) and never cooperate with a one-shot belief (*Cooperation Probability$_{One-Shot}$* ≈ 0) (*SI Text*). If the canonically rational decision rules were in fact optimal, then selection should not change these values.

We ran 3,000 simulations in this set (*SI Text*). Fig. 3A shows how *Cooperation Probability$_{One-Shot}$* evolves in several typical runs. For these examples, half of the dyads are one-shot, and the distance between the cue summary distributions is two (the middle value of the distances that we examined). The benefit to cost ratio within a PD round is 3:1, a relatively small value. Given these parameters, if an agent believes that she is facing a one-shot PD, there is only a 16% chance that she is wrong and instead faces a repeated PD. Nonetheless, when the average length of a repeated interaction is 5 or 10 rounds—relatively short interaction lengths—agents with a one-shot belief nonetheless evolve to cooperate a remarkable 87% or 96% of the time, respectively. When discrimination is even easier—when the distance between distributions is three—there is only a 7% chance that an agent with a one-shot belief is actually facing a repeated interaction. In this case, when the average number of interactions is only 10, then agents with a one-shot belief evolve to cooperate 47% of the time, despite the fact that such interactions will turn out to be repeated only 7% of the time (*SI Text*).

Fig. 3B summarizes the values that *Cooperation Probability$_{One-Shot}$* evolves to as a function of (*i*) the benefits that can be gained in a single round of cooperation and (*ii*) the average length of repeated interactions, with the other parameters fixed. As predicted, *Cooperation Probability$_{One-Shot}$* evolves to higher values when either of these variables increases. When benefit size and interaction length parameters exceed those used in the examples above, *Cooperation Probability$_{One-Shot}$* evolves to be extremely high. In many cases, the regulatory variables evolve such that agents almost always cooperate, even with an explicit belief that their interaction is one-shot (Fig. 3B, Figs. S1 and S2, and Table S1).

Simulation Set 2. Given behavior that is consistent with belief, how will selection shape the formation of such beliefs? We conducted an additional set of simulations to explore this case, where the cognitive rule, but not the motivational rule, was allowed to evolve. The motivational rule was fixed at what is conventionally considered to be optimal: *Defect if you believe that the interaction is one-shot; Otherwise, cooperate.* Here, agents do not form optimally rational beliefs by Bayesian updating. Instead, agents execute a fast and frugal heuristic (42); agents simply observe their partner's cue summary, compare it with an evolvable decision threshold value, and believe the interaction is repeated or not based on that comparison. This threshold value can be seen as the level of evidence that the agent requires before it believes that an interaction is repeated and chooses to cooperate. As before, to work against our hypothesis, we set the average threshold values for first-generation agents at the value that equalized the rates of the two error types (e.g., a threshold of 0 when the base rates of one-shot and repeated dyads are both 0.5).

Paralleling the previous results, here, agents also typically evolve to cooperate, despite strong evidence that the interaction is one-shot. For example, with the same parameter values used in the upper lines of Fig. 3A (i.e., where cooperation has only modest benefits), Fig. 3C shows that agents evolve to have belief thresholds that lead them to cooperate ~60% of the time when their interaction is actually one-shot. In contrast, given these thresholds, agents will only defect 1% of the time when their interaction is actually repeated.

Again, paralleling the previous results, Figs. 3D, Figs. S3 and S4 and Table S2 demonstrate that, across multiple simulation runs, greater within-round benefits to cooperation and longer interactions cause the evolution of agents who are increasingly likely to cooperate, even when given evidence that they are in a one-shot interaction. This move to increasing cooperativeness is accomplished by creating a cognitive rule that requires very high levels of evidence before it will conclude that the interaction is one-shot. As expected, the cognitive architecture evolves to make fewer expensive errors (defecting in repeated interactions) at the cost of a higher frequency of the cheaper errors (cooperating in one-shot interactions). Cognitive architectures that are highly resistant to concluding that interactions are one-shot are favored by selection over architectures that are cognitively more accurate.

It is important to ensure that these two sets of results are not fragilely predicated on specific assumptions of our formalizations. We checked the robustness of these two parallel results in several ways. First, we reran both sets of simulations but allowed agents to erroneously defect when their strategy specifies cooperation on a given round. Errors have been shown to have important consequences for evolutionary dynamics, revealing hidden assumptions in models that, once removed, can prove evolutionarily fatal to certain strategies (43). In these simulations, we used the GRIM strategy, because it militates against our hypothesis. Although including such errors necessarily lowers the average benefits of repeated cooperation—because the GRIM strategy shuts down cooperation after a single experience of defection—the qualitative dynamics of our simulations remain unchanged. Agents still robustly evolve to cooperate in one-shot interactions (*SI Text* and Figs. S2 and S4).
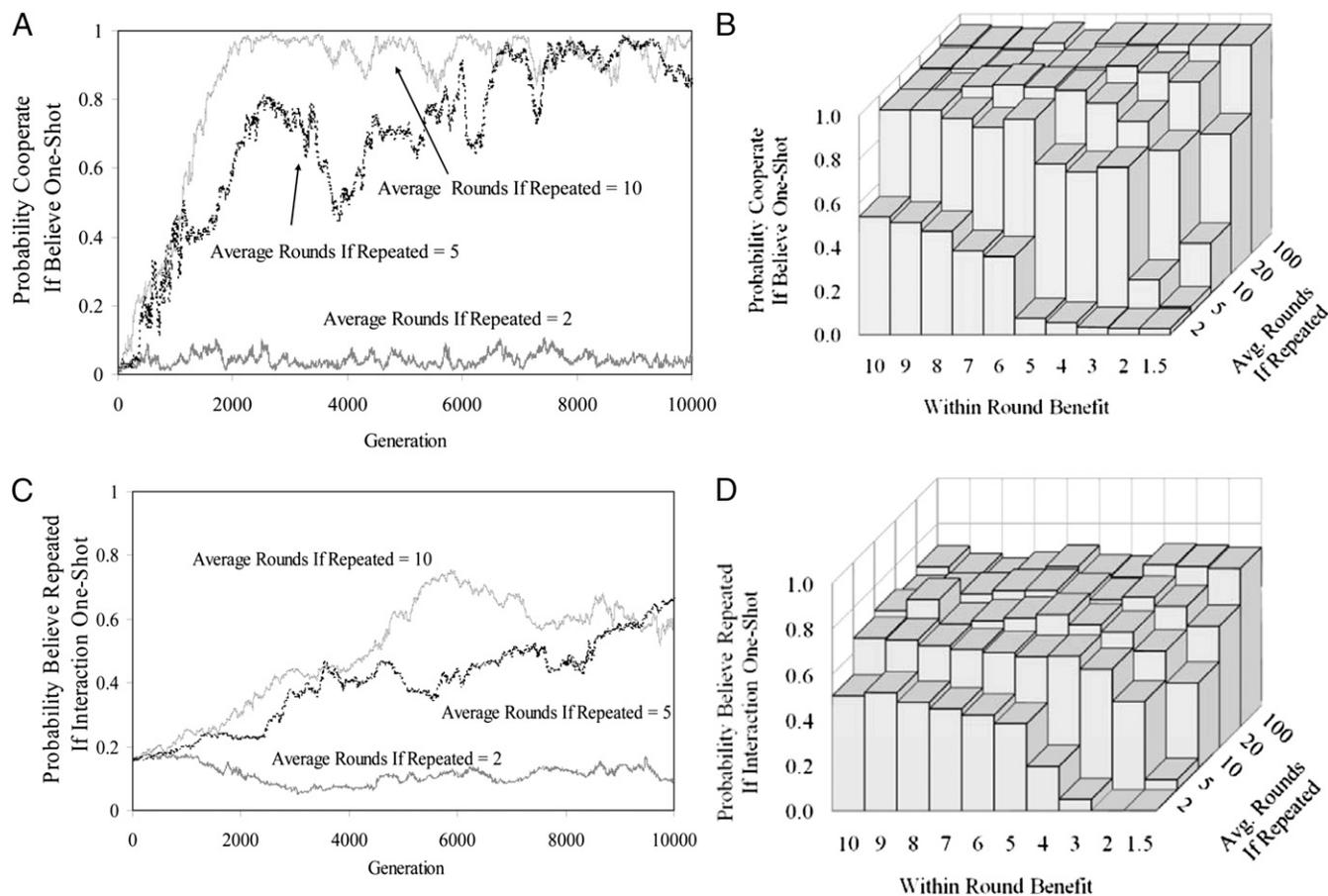
Fig. 3. One-shot cooperation evolves when interactions are moderately long and benefits are moderately large. (*A*) Example evolutionary dynamics showing that the probability of cooperation, given a rational (ideal Bayesian) belief that an interaction is one-shot, evolves to be high when the motivational architecture is allowed to evolve based on its fitness consequences. (*B*) Aggregated final stabilized values across all runs where the motivational architecture is free to evolve and half of all interactions are repeated; this example shows that the probability of cooperating despite a rational belief that an interaction is one-shot evolves to be high except when the benefits of cooperating and the average length of repeated interactions are both small. (*C*) Example evolutionary dynamics showing that agents evolve to become highly resistant to concluding that interactions are one-shot, even in the face of strong evidence that interactions are one-shot, when the weight of evidence required by the cognitive architecture is allowed to evolve. These stringent thresholds to conclude the interaction is one-shot results in a high probability of cooperating, even when the interaction is actually one-shot. (*D*) Aggregated final stabilized values across all runs where the cognitive architecture is free to evolve and one half of all interactions are repeated. These data show that the probability of cooperating when the interaction is one-shot evolves to be high, except when benefits of cooperating and the average length of repeated interaction are both small. For *A* and *C*, the benefit to cost ratio is 3:1, half of all interactions are repeated, and the distance between cue summary distributions (a measure of the ease of discriminating one-shot from repeated interactions) is 2 SDs. Traditionally, it has been assumed that the highest paying strategy is to cooperate 0% of the time whenever an interaction is likely to be one-shot.

As a final check, we created a simple, nondynamic, best-response analytic model of ideal behavior in this kind of uncertain ecology. Paralleling the simulation results, this model shows that agents will be designed to cooperate when they believe that the interaction is one-shot if interactions are sufficiently long and within-round benefits sufficiently large. Indeed, when the benefits of repeated cooperation are sufficiently large, agents should always cooperate, even when presented evidence that they face a one-shot interaction (*SI Text*, Eqns. S1 and S2). Regardless of the method used, cooperation in one-shot encounters is a robust result of selection for direct reciprocity, once the necessary step of discriminating one-shot from repeat encounters is explicitly included as part of the decision problem.

Discussion

Despite the fact that cooperation in one-shot interactions is viewed as both biologically maladaptive and economically irrational, it is nonetheless behaviorally widespread in our species. This apparent anomaly has posed a challenge to well-established theories in biology and economics, and it has motivated the development of a diverse array of alternatives—alternatives that seem to either conflict with known selection pressures or sensitively depend on extensive sets of untested assumptions.

These alternatives all assume that one-shot cooperation is an anomaly that cannot be explained by the existence of cooperative architectures that evolved for direct reciprocity. Our main results show that this assumption is false: organisms undergoing nothing but a selective regime for direct reciprocity typically evolved to cooperate even in the presence of strong evidence that they were in one-shot interactions. Indeed, our simulated organisms can form explicit beliefs that their interactions are one-shot and, nonetheless, be very likely to cooperate. By explicitly modeling the informational ecology of cooperation, the decision-making steps involved in operating in this ecology, and selection for efficiently balancing the asymmetric costs of different decision errors, we show that one-shot cooperation is the expected expression of evolutionarily well-engineered decision-making circuitry specialized for effective reciprocity.

This cooperation-elevating effect is strong across broad regions of parameter space. Although it is difficult to precisely

map parameters in simplified models to real-world conditions, we suspect that selection producing one-shot generosity is likely to be especially strong for our species. The human social world—ancestrally and currently—involves an abundance of high-iteration repeat interactions and high-benefit exchanges. Indeed, when repeated interactions are at least moderately long, even modest returns to cooperation seem to select for decision architectures designed to cooperate even when they believe that their interaction will be one-shot. We think that this effect would be even stronger had our model included the effects of forming reputations among third parties. If defection damages one's reputation among third parties, thereby precluding cooperation with others aside from one's current partner, defection would be selected against far more strongly (44). Therefore, it is noteworthy that cooperation given a one-shot belief evolves even in the simple case where selection for reputation enhancement cannot help it along. It is also worth noting that a related selection pressure—defecting when you believe your partner will not observe you—should be subject to analogous selection pressures. Uncertainty and error attach to judgments that one's actions will not be observed, and the asymmetric consequences of false positives and misses should shape the attractiveness of defection in this domain as well.

In short, the conditions that promote the evolution of reciprocity—numerous repeat interactions and high-benefit exchanges—tend to promote one-shot generosity as well. Consequently, one-shot generosity should commonly coevolve with reciprocity. This statement is not a claim that direct reciprocity is the only force shaping human cooperation—only that if reciprocity is selected for (as it obviously was in humans), its existence casts a halo of generosity across a broad variety of circumstances.

According to this analysis, generosity evolves because, at the ultimate level, it is a high-return cooperative strategy. Yet to implement this strategy at the proximate level, motivational and representational systems may have been selected to cause generosity even in the absence of any apparent potential for gain. Human generosity, far from being a thin veneer of cultural conditioning atop a Machiavellian core, may turn out to be a bedrock feature of human nature.

1. Trivers RL (1971) The evolution of reciprocal altruism. *Q Rev Biol* 46:35–57.
2. Axelrod R, Hamilton WD (1981) The evolution of cooperation. *Science* 211:1390–1396.
3. Hauert C, Michor F, Nowak MA, Doebeli M (2006) Synergy and discounting of cooperation in social dilemmas. *J Theor Biol* 239:195–202.
4. Hammerstein P (2003) *The Genetic and Cultural Evolution of Cooperation* (MIT Press, Cambridge, MA).
5. Maynard Smith J (1982) *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, UK).
6. Cosmides L, Tooby J (2005) Neurocognitive adaptations designed for social exchange. *The Handbook of Evolutionary Psychology*, ed Buss DM (Wiley, New York), pp 584–627.
7. Camerer CF (2003) *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton University Press, Princeton).
8. Ermer E, Guerin SA, Cosmides L, Tooby J, Miller MB (2006) Theory of mind broad and narrow: Reasoning about social exchange engages ToM areas, precautionary reasoning does not. *Soc Neurosci* 1:196–219.
9. Sugiyama LS, Tooby J, Cosmides L (2002) Cross-cultural evidence of cognitive adaptations for social exchange among the Shiwiar of Ecuadorian Amazonia. *Proc Natl Acad Sci USA* 99:11537–11542.
10. Krueger F, et al. (2007) Neural correlates of trust. *Proc Natl Acad Sci USA* 104: 20084–20089.
11. McCabe K, Houser D, Ryan L, Smith V, Trouard T (2001) A functional imaging study of cooperation in two-person reciprocal exchange. *Proc Natl Acad Sci USA* 98: 11832–11835.
12. Stone VE, Cosmides L, Tooby J, Kroll N, Knight RT (2002) Selective impairment of reasoning about social exchange in a patient with bilateral limbic system damage. *Proc Natl Acad Sci USA* 99:11531–11536.
13. Yamagishi T, Terai S, Kiyonari T, Mifune N, Kanazawa S (2007) The social exchange heuristic: Managing errors in social exchange. *Ration Soc* 19:259–291.
14. de Quervain DJF, et al. (2004) The neural basis of altruistic punishment. *Science* 305: 1254–1258.
15. Knoch D, Pascual-Leone A, Meyer K, Treyer V, Fehr E (2006) Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314:829–832.
16. Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003) The neural basis of economic decision-making in the Ultimatum Game. *Science* 300:1755–1758.
17. Kelly RL (1995) *The Foraging Spectrum* (Smithsonian Institution Press, Washington, DC).
18. Dawes RM, Thaler RH (1988) Anomalies: cooperation. *J Econ Perspect* 2:187–197.
19. Fehr E, Fischbacher U, Gachter S (2002) Strong reciprocity, human cooperation, and the enforcement of social norms. *Hum Nat* 13:1–25.
20. Henrich J, et al. (2005) "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behav Brain Sci* 28:795–815.
21. McCabe KA, Rigdon ML, Smith VL (2003) Positive reciprocity and intentions in trust games. *J Econ Behav Organ* 52:267–275.
22. Fehr E, Henrich J (2003) Is strong reciprocity a maladaptation? On the evolutionary foundations of human altruism. *Genetic and Cultural Evolution of Cooperation*, ed Hammerstein P (MIT Press, Cambridge, MA), pp 55–82.
23. Gibbons R (1992) *Game Theory for Applied Economists* (Princeton University Press, Princeton).
24. Boyd R, Gintis H, Bowles S, Richerson PJ (2003) The evolution of altruistic punishment. *Proc Natl Acad Sci USA* 100:3531–3535.
25. Fehr E, Fischbacher U (2003) The nature of human altruism. *Nature* 425:785–791.
26. Van Vugt M, Van Lange PAM (2006) The altruism puzzle: Psychological adaptations for prosocial behavior. *Evolution and Social Psychology*, eds Schaller M, Simpson JA, Kenrick DT (Psychosocial Press, Madison, CT), pp 237–261.
27. Henrich J (2004) Cultural group selection, coevolutionary processes and large-scale cooperation. *J Econ Behav Organ* 53:3–35.
28. Gintis H (2000) Strong reciprocity and human sociality. *J Theor Biol* 206:169–179.
29. Haidt J (2007) The new synthesis in moral psychology. *Science* 316:998–1002.
30. Wilson DS, Sober E (1994) Re-introducing group selection to the human behavioral sciences. *Behav Brain Sci* 17:585–654.
31. Burnham TC, Johnson DDP (2005) The biological and evolutionary logic of human cooperation. *Anal Kritik* 27:113–135.
32. Delton AW, Krasnow MM, Cosmides L, Tooby J (2010) Evolution of fairness: Rereading the data. *Science* 329:389.
33. Hagen EH, Hammerstein P (2006) Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theor Popul Biol* 69:339–348.
34. Lehmann L, Rousset F, Roze D, Keller L (2007) Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. *Am Nat* 170: 21–36.
35. Trivers R (2004) Genetic and cultural evolution of cooperation. *Science* 304:964–965.
36. West SA, Griffin AS, Gardner A (2007) Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection. *J Evol Biol* 20:415–432.
37. Cosmides L (1989) The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition* 31:187–276.
38. Green DM, Swets JA (1966) *Signal Detection Theory and Psychophysics* (Wiley, New York).
39. Haselton MG, Nettle D (2006) The paranoid optimist: An integrative evolutionary model of cognitive biases. *Pers Soc Psychol Rev* 10:47–66.
40. Kiyonari T, Tanida S, Yamagishi T (2000) Social exchange and reciprocity: Confusion or a heuristic? *Evol Hum Behav* 21:411–427.
41. Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND (2011) How to grow a mind: Statistics, structure, and abstraction. *Science* 331:1279–1285.
42. Todd PM, Gigerenzer G (2000) Précis of Simple heuristics that make us smart. *Behav Brain Sci* 23:727–741.
43. Panchanathan K, Boyd R (2003) A tale of two defectors: The importance of standing for evolution of indirect reciprocity. *J Theor Biol* 224:115–126.
44. Nowak MA, Sigmund K (2005) Evolution of indirect reciprocity. *Nature* 437: 1291–1298.

Delton et al.

Supporting Information

Delton et al. 10.1073/pnas.1102131108

SI Text

In the main text, we give examples showing that organisms undergoing nothing but a selective regime for direct reciprocity evolve to cooperate even in the presence of strong evidence that they are in one-shot interactions. Here, we detail the simulations used to generate these results, give a complete summary of the results, and develop an analytic model that complements the results. The main result to note is that, across very broad ranges of parameter space, cooperation in one-shot interactions is a robust outcome. It is also worth noting that the relatively small regions in which one-shot cooperation does not evolve heavily overlap with the regions where cooperation itself will not evolve, because repeat interactions are too few in relation to the benefits of cooperation for cooperation to be evolutionarily favored.

Simulation Details. To assess the evolutionary dynamics of one-shot encounters when interactions can either be one-shot or repeated, we conducted a series of agent-based simulations. Each simulation run consisted of a population of 500 agents allowed to evolve through 10,000 generations. For each generation, the population is randomly sorted, irrespective of genotype, into dyads. Dyads are then randomly assigned, irrespective of genotype, to interact in one-shot or indefinitely repeated prisoner's dilemmas (PDs). The base rate of one-shot dyads was varied across simulation runs at 10%, 30%, 50%, 70%, and 90%. These base rates for one-shot interactions were selected because they are high and, therefore, work against our hypothesis. Indeed, for our band-living hunter–gatherer ancestors, even 10% is likely an order of magnitude too high. These values serve as worst-case scenarios; the greater the base rate of one-shot interactions, the more decision architectures will be selected to respond adaptively to them, creating conditions more likely to favor defection.

Each member of a dyad independently (and randomly with respect to their own genotype) draws a cue summary from the appropriate distribution (Fig. 2). That is, each agent assigned to a one-shot interaction dyad independently draws a cue summary from the distribution for one-shot interactions, and each agent assigned to a repeated interaction dyad independently draws a cue summary from the distribution for repeated interactions.

In both types of dyads, the interaction consists of at least one round of interaction. If the interaction was one-shot, there was no additional interaction. If the interaction was indefinitely repeated, the two agents moved onto another round of interaction with probability w, with values of w varying across simulation runs from 0.5, 0.8, 0.9, 0.95 to 0.99. Whether a dyad moves from (e.g.) round 1 to round 2 is computed independently from whether it moves from (e.g.) round 2 to round 3. Thus, the average number of rounds for a repeated dyad is $1/(1 − w)$, with values varying across simulation runs from 2, 5, 10, 20 to 100 average rounds of interaction in repeated dyads. Although we contrast indefinitely repeated interactions with one-shot interactions, because the latter are the simplest and most extreme cases, what is really at issue is the distinction between interactions with indefinite endpoints and interactions with finite endpoints. In the latter situation, the payoff-maximizing choice is for agents to defect in every round of interaction (1). Of all interactions with finite endpoints, one-shot interactions have the least potential to be profitable (e.g., they have the lowest potential for generating gains through mistaken cycles of mutual cooperation). Using one-shot encounters is, therefore, the most difficult test case for the evolution of cooperation in finite interactions.

During each round of their PD, each agent can give a benefit b to the other agent at a cost c to the self. There are gains in trade such that $b > c > 0$. For the simulations, we hold c constant at one, meaning b also represents the benefit to cost ratio. The fitness that an agent earns in each interaction is added to a constant baseline fitness. Because we are interested in the effects of selection, we set baseline fitness to 10 to ensure that, given the strategies and parameters used, no agent's fitness will ever be less than 1. The values of b varied across simulation runs from 1.5, 2, 3, 4, 5, 6, 7, 8, 9 to 10.

Based on their decision rules and the cue summary that they observe from their partner (both described below), agents either play a strategy of always defecting (ALLD) or the well-known TIT-for-TAT strategy (TFT) (2, 3). TFT cooperates on the first round of interaction and thereafter only cooperates when its partner cooperated in the previous round; otherwise, TFT defects. Note that our goal is to model a choice between cooperation and defection, not to determine exactly which cooperative strategy would perform best in the world modeled. There are many variants and alternatives to TFT, many of which outperform it. However, at least when playing themselves, all of these variants are fundamentally cooperative strategies. Compared with the alternatives, TFT has the additional benefit of being familiar to most readers. (Indeed, our own view is that the evolved human cooperative psychology plays a strategy or metastrategy that is far more complex than existing formalized cooperative strategies; our selection of a simple cooperative strategy is merely illustrative.)

After being sorted into one-shot or repeated dyads, agents are assigned a cue summary that is probabilistically associated with the type of dyad the agent belongs to. These cue summaries are drawn independently for both agents in a dyad and are drawn independently of the organisms' genotype (i.e., independently of their heritable strategy). In real life, a variety of cues could be probabilistically associated with whether the dyad is one-shot or repeated, such as prior history of interaction, community of residence, etc. As for many phenomena that involve summing multiple error-prone quantities, cue summaries are modeled as a single normal distribution.

Agents can see their partner's cue summary but cannot see their own. Each agent's cue summary is independently drawn from a normal distribution with a SD of 1 and a mean of $−d/2$ or $d/2$ for the one-shot and repeated dyads, respectively. Cue summaries discriminate between the two types of dyads probabilistically (rather than perfectly); that is, there is some overlap between the distribution of cue summaries associated with one-shot interactions and the distribution associated with repeated interactions. The ease of discrimination is defined by the parameter d (the distance between the means of the two distributions), which took on the values one, two, and three—corresponding to a proportional overlap between the distributions of 61.7%, 31.7%, and 13.4%, respectively.

We conducted two simulation sets, each with a different cooperative rule. One simulation set is designed to test the evolutionary dynamics of decision-making components concerned with representing states of the world (cognitive components). The other simulation set is designed to test the evolutionary dynamics of decision-making components that transform such representations into decisions about what actions to take (motivational components).

The first simulation set (SS1) investigates the evolutionary dynamics of motivational components. In these simulations, the

agents form beliefs using Bayesian updating—allowing them to have the most accurate beliefs possible—but then, translate those beliefs into behavior using evolvable motivational variables. In this simulation set, it is assumed that the agents have innate and perfect knowledge of the actual base rates of one-shot and repeated interactions in their world. Based on Bayesian updating, if the posterior probability that the dyad is one-shot is largest, then the agent believes that the dyad is one-shot. The agent then generates a random number on the uniform interval [0, 1] and compares that with an internal regulatory variable that represents the probability of cooperating given that the agent believes that the dyad is one-shot, $Probability\ Cooperation_{One\text{-}Shot}$. If the random number is less than $Probability\ Cooperation_{One\text{-}Shot}$, the agent cooperates by playing TFT; otherwise, the agent defects. If the posterior probability that the dyad is repeated is largest, the agent goes through a similar algorithm. In this case, the agent consults an internal regulatory variable that represents the probability of cooperating by playing TFT given a belief that the interaction is repeated, $Probability\ Cooperation_{Repeated}$. Over generations, the magnitudes of these internal regulatory variables are allowed to evolve. To work against our hypothesis, at the start of each simulation run, we set these evolvable variables at values that maximize consistency with belief: $Probability\ Cooperation_{One\text{-}Shot} \approx 0$ and $Probability\ Cooperation_{Repeated} \approx 1$ (see paragraph on mutation below).

Our second simulation set (SS2) investigates the evolutionary dynamics of cognitive components. In these simulations, we fix an agent's actions to be consistent with its belief—cooperating (by playing TFT) in repeated interactions and, otherwise, defecting. However, beliefs are not computed with Bayesian updating. Instead, agents compare their partner's cue summary—their only evidence about whether the interaction is one-shot or repeated—with an evolvable internal threshold to determine their belief. This threshold can take on the same range of values as the cue summary: any positive or negative number. If the partner's cue summary is greater than the threshold, then the agent believes that the interaction will be repeated (and therefore, cooperates by playing TFT); otherwise, the agent believes that the interaction is one-shot (and therefore, defects). The value of the threshold reflects an agent's default assumption about its interactions. As agents evolve increasingly positive thresholds, they demand more and more evidence that an interaction will be repeated before they believe that it is repeated and, therefore, cooperate (i.e., they require increasingly higher cue summaries). Conversely, as agents evolve increasingly negative thresholds, they become more credulous; they require less evidence before believing that they are having a repeated interaction (i.e., even low cue summaries will trigger the belief that the interaction is repeated). Again, to advantage the alternative hypothesis, we start the first generation with an average threshold that maximizes accuracy (see paragraph on mutation below).

For every simulation run, the population evolves through 10,000 generations. The probability that an agent in the current generation is the parent of a given member of the next generation is proportional to the agent's fitness. Specifically, the probability that an agent was the parent of a given member of the next generation was the ratio of its fitness to the total fitness of its entire generation. Algorithmically, agents were assigned intervals of probability space based on their proportion of the population fitness. Offspring were created by randomly sampling from this probability space until a complete successor generation was filled. Fitness only probabilistically influenced an agent's number of offspring, allowing the possibility of genetic drift. Offspring inherited their parent's genotype with a 5% probability of mutation, and, when mutations occurred, they modified the parent's genotype by a normally distributed random variable with a mean = 0 and SD = 0.025. For the first simulation set, mutations are independent for the two regulatory variables. How-

ever, because they represent probabilities, mutations are not allowed to move these values to greater than one or less than zero. To generate variance on which natural selection can act, in the initial generation of both simulation types, all agents are run through the mutation procedure described above, except that there is a 100% chance that a mutation takes place.

We factorially explored the parameter space of b, w, d, and the base rate of one-shot dyads; these parameters had 10, 5, 3, and 5 possible values, respectively. Four independent simulation runs were conducted for each instantiation of the parameters. Therefore, for each simulation type, there were $10 \times 5 \times 3 \times 5 \times 4 = 3{,}000$ runs. There were two types of decision rules, creating 6,000 runs. We then replicated this procedure introducing errors in agents' behavior (see below) for a total of 12,000 simulation runs.

Theoretical work has shown errors to have an important impact on the evolutionary dynamics of games by exposing hidden weaknesses of strategies, potentially destabilizing and leading to the extinction of nonrobust strategies (4). Therefore, we conducted a second set of simulations identical to those simulations described above, except that we incorporated errors. When an error occurred, an agent who intended to cooperate instead defected. The reverse was not true; agents never mistakenly cooperated when they intended to defect. The error rate was set at a relatively high proportion: 0.05. In other words, in 1 of 20 instances where an agent intends to cooperate, it defects instead. In these simulations, agents use GRIM, a variant of TFT. GRIM cooperates until it observes that its partner defected. After GRIM defects, it defects forever. When there are no errors, GRIM and TFT behave identically when playing themselves or ALLD. When there are errors, GRIM fails to effectively cooperate in situations where TFT would, thus lowering the average benefit of playing a cooperative strategy. Because our hypothesis depends on successful cycles of cooperation causing large benefits, using GRIM works against our hypothesis.

Simulation Results. The full results of the main simulations (i.e., those without errors) are presented in Figs. S1 and S3. These figures show how selection exclusively based on payoffs to direct reciprocity crafts organisms to be "irrationally" cooperative. As predicted, cooperation in one-shot encounters becomes increasingly likely as the within-round benefits of cooperation increase and the expected length of interactions increases. The inclusion of errors does not drastically change the evolutionary dynamics, although it does mitigate the magnitude of the effects (Figs. S2 and S4). In all cases, the results were computed by averaging the final 500 generations of the individual simulations. Because the evolvable thresholds from the second simulation set are not in an easily interpretable metric, for Figs. S3 and S4, we converted them to the expected probability that an agent with that threshold would believe the interaction is repeated and, therefore, would cooperate if it was in fact facing a one-shot interaction. This conversion is accomplished by simply determining the value at the threshold of a cumulative distribution function for a normal distribution with a mean = $-d/2$ and an SD = 1. Additionally, the main data were analyzed as multiple regressions to provide quantified estimates of the main effects (Tables S1 and S2).

As one might expect, greater ease of discrimination (i.e., higher values of d) selected for agents who were less willing to cooperate in one-shot interactions. However, the difficulty of discriminating one-shot from repeated dyads did not generally have a large effect on the evolution of regulatory variables: the amount of variance explained by d was 1.1% and 0.5% in SS1 and SS2, respectively, and was far less than that explained by each of the other variables, which ranged from 2% to 33% (because the four predictors are perfectly orthogonal, percentage of variance explained can be calculated by squaring the β of each variable). Thus, our summary graphs average over this parameter.

Finally, note that there are, nevertheless, some regions of parameter space where one-shot cooperation does not evolve. However, these regions are closely associated with areas where cooperation in repeated interactions will not evolve. Reciprocity will be selected against unless $w \times b > c$. Hence, even in a world where interactions are guaranteed the chance of being repeated, if the average length of repeated interactions is 2 ($w = 0.5$) and benefits are <2, then cooperation is not favored. Therefore, it is to be expected that these conditions, which select against cooperation in repeated interactions, also select against one-shot cooperation.

Strengths and limitations of each model. Each simulation set has different strengths and limitations. Many economists prefer models in which preferences (motivations) vary, but beliefs are formed rationally, as in SS1. By contrast, many cognitive psychologists doubt that beliefs are rationally formed using Bayesian updating and point to heuristics and biases that affect the judgments people make—a view more consistent with SS2. Evidence consistent with either model can be found depending on which area of the literature that a reader consults.

For example, some economists may view the evidence from experimental games as more consistent with SS1, given that (*i*) the experimenter informs subjects that the interaction was one-shot (or repeated, as the case may be) and (*ii*) most subjects verbally report that they believe what the experimenter told them (e.g., 5, 6). For the sake of argument, let us assume that these subjects are trying to report their beliefs truthfully (and not just being polite by telling the experimenter that they believe what he or she told them). These verbal reports, albeit face valid, would not settle the SS1 vs. SS2 issue for most psychologists because research in cognitive neuroscience (e.g., split brain patients), psychology (social, cognitive, developmental, and evolutionary), and even behavioral economics shows that behavior can be regulated by nonconscious or implicit variables that reflect assessments of the situation (rather than preferences) and that the content of verbal reports is often uncorrelated with these variables.

A verbal report that "X is true" can be associated with different degrees of confidence. What you say is true can be different from your gut feeling about what is true (a decision threshold), with consequential behaviors regulated by the decision threshold (as in SS2) rather than by a verbal statement (that was elicited by an experimenter's query). Indeed, given that speaking evolved in the context of navigating a complex social world, decisions about what to say, admit to, or advocate may be generated by different cognitive systems than decisions about what actions to take.

Psychologists—and many economists—would also point out that situational cues relevant to epistemic judgments about facts of the matter (is this situation one-shot? is it really anonymous?) importantly affect behavior in experimental games, even when these cues conflict with what the experimenter told the subjects (and with what the subjects report that they were told). For example, economists often find that verbal assurances of anonymity are not sufficient to elicit more selfish decisions in economic games, yet, generosity does decrease when experimenters create a transparent situation that guarantees anonymity (7) and it increases in the presence of stylized eyes—an ancestrally reliable cue that one is being observed (8). These situational cues could exert their effect by Bayesian updating (as in SS1) or by shifting the weight of evidence above a decision threshold (as in SS2). Whichever is true, the subject might report that they believe the situation to be anonymous—not because they are lying or being overly polite but because the system that generates these verbal judgments does not have any more access to the computations that regulate cooperative behavior than it does to the retinal smears and low-level line detectors that build our conscious (and verbally reportable) representation of the visual world.

So, should one prefer SS1 or SS2? Our own view is that natural selection has probably tinkered with both motivational and cognitive components of the architecture that regulate cooperative behavior. But, rather than prejudge the issue, we chose to model each component separately while holding the other constant. The results show that a disposition for generosity evolves in either case; future research will have to determine the precise design of the computational systems that actually evolved.

Analytic Model. Here, we develop a simple, nonevoluntionary, best-response model showing that cooperating even in the face of an explicit one-shot belief is always the payoff-maximizing choice when the long-term net benefits of cooperation are large enough. This analytic model makes the same assumptions as the model of the main paper:

i) Interactions are structured as PDs such that, on each round of an interaction, each member of a dyad can choose whether to transfer a benefit b to their partner at a personal cost c. There are gains in trade (i.e., $b > c$).

ii) Some interactions are one-shot; this type of interaction occurs with probability P. Other interactions are repeated; this type of interaction occurs with probability $1 - P$.

iii) However, organisms can only probabilistically determine whether their own interaction is one-shot or repeated by observing their partner's cue summary. Cue summaries are drawn from two standard normal distributions that differ in mean value depending on whether the interaction is one-shot or repeated.

iv) Organisms have perfect, innate knowledge of the cue summary distributions, including their base rates (symbolized by P) and their relative distances from each other.

v) Given their partner's cue summary and this knowledge of the cue summary distributions, organisms use Bayesian updating to compute posterior probabilities of their current interaction being one-shot or repeated.

We can avoid representing the process of Bayesian updating explicitly and thus drastically simplify the model. Given that the parameters underlying the distribution of the cue summaries are fixed, then there are also fixed probabilities that an organism will come to believe its interaction is one-shot. We denote the probability that an organism comes to believe its interaction is one-shot when the interaction is, in fact, one-shot as Θ_{OS}, and we denote the probability that an organism comes to believe its interaction is one-shot when the interaction is, in fact, repeated as Θ_R. The complements of these probabilities denote the probabilities that an organism comes to believe its interaction is repeated. If an organism believes its interaction is one-shot, it cooperates with probability α_i, where the subscript i (equal to one or two) denotes which member of the dyad this value belongs to. If an organism believes that its interaction is repeated, it cooperates with probability β_i. These final two variables were previously referred to as *Probability Cooperation$_{One\text{-}Shot}$* and *Probability Cooperation$_{Repeated}$*. However, given that here they are embedded within equations, we use a more concise notation.

Given this specification, we derive payoff functions for the organisms. As an example of part of the payoff function, the expression $(1 - P)(1 - \Theta_R)\beta_1\Theta_R\alpha_2((b - c)/(1 - w))$ represents, going from left to right, the probability that the interaction is in fact repeated, multiplied by the probability that player 1 believes the interaction is repeated (given that it is in fact repeated), multiplied by the probability that player 1 cooperates given a repeated belief, multiplied by the probability that player 2 believes the interaction is one-shot (given that it is actually repeated), multiplied by the probability player 2 cooperates given a one-shot belief, multiplied by the final quantity enclosed in parentheses, which represent the net benefits of sustained mu-

tual cooperation. The full payoff function for player 1, V_1, is given as Eq. S1 (because players are symmetrical, this equation also applies to player 2, with appropriate subscript substitutions).

$$V_1 = P\left[\Theta_{OS}\begin{bmatrix}\Theta_{OS}(\alpha_1\alpha_2(b-c) - \alpha_1(1-\alpha_2)c + (1-\alpha_1)\alpha_2 b)\\ + (1-\Theta_{OS})(\alpha_1\beta_2(b-c) - \alpha_1(1-\beta_2)c + (1-\alpha_1)\beta_2 b)\end{bmatrix} + (1-\Theta_{OS})\begin{bmatrix}\Theta_{OS}(\beta_1\alpha_2(b-c) - \beta_1(1-\alpha_2)c + (1-\beta_1)\alpha_2 b)\\ + (1-\Theta_{OS})(\beta_1\beta_2(b-c) - \beta_1(1-\beta_2)c + (1-\beta_1)\beta_2 b)\end{bmatrix}\right] + (1-P)\left[\Theta_R\begin{bmatrix}\Theta_R\left(\alpha_1\alpha_2\dfrac{b-c}{1-w} - \alpha_1(1-\alpha_2)c + (1-\alpha_1)\alpha_2 b\right)\\ + (1-\Theta_R)\left(\alpha_1\beta_2\dfrac{b-c}{1-w} - \alpha_1(1-\beta_2)c + (1-\alpha_1)\beta_2 b\right)\end{bmatrix} + (1-\Theta_R)\begin{bmatrix}\Theta_R\left(\beta_1\alpha_2\dfrac{b-c}{1-w} - \beta_1(1-\alpha_2)c + (1-\beta_1)\alpha_2 b\right)\\ + (1-\Theta_R)\left(\beta_1\beta_2\dfrac{b-c}{1-w} - \beta_1(1-\beta_2)c + (1-\beta_1)\beta_2 b\right)\end{bmatrix}\right]$$

[S1]

To find the best response of player 1, we take the derivative of this function with respect to α_1. This quantity, with some algebraic rearrangement, is shown as Eq. S2.

$$\frac{b-c}{(1/w)-1}(1-P)\Theta_R(\Theta_R\alpha_2 + (1-\Theta_R)\beta_2) - c(P\Theta_{OS} + (1-P)\Theta_R)$$

[S2]

This expression no longer contains any terms involving α_1. Therefore, we need simply to determine whether the total quantity is greater than or less than zero. If the derivative is less than zero, then the best response is to never cooperate with a one-shot belief. However, if the derivative is greater than zero, then the best response is to always cooperate with a one-shot belief. The total left-hand quantity represents the net payoffs of repeated cooperation (albeit discounted by an additional factor of w) multiplied by the probability that the focal agent's partner cooperates when the interaction is in fact repeated and the probability that the focal agent believes the interaction is one-shot. The total right-hand quantity represents the cost of co-operation when the partner does not cooperate multiplied by the probability of an agent believing that the interaction is one-shot. All parameters denoting probability (with the exception of α_2) can reasonably be assumed to be strictly greater than zero: We assume that either type of interaction can actually occur. Under this assumption, the probabilities of believing that an interaction is one-shot are necessarily strictly greater than zero. Given that, outside a narrow range of values cooperation in repeated interactions is favored, the probability of cooperating given a belief that the interaction is repeated should also be strictly greater than zero, possibly near one. At the very least, given the assumptions of the model, the range of parameters leading to nonzero probabilities of cooperation are less restrictive for β than for α. Given these assumptions, the total expression will be greater than zero as long as the benefits of cooperation are sufficiently large compared with the costs and as long as repeated interactions are sufficiently long.

This analysis confirms the reasoning and simulations already presented: When there are sizeable gains in trade to be made through mutually beneficial, long-lasting exchange, organisms should cooperate in one-shot encounters even when they have explicit beliefs that cooperation is one-shot.

1. Gibbons R (1992) *Game Theory for Applied Economists* (Princeton University Press, Princeton).
2. Axelrod R, Hamilton WD (1981) The evolution of cooperation. *Science* 211:1390–1396.
3. Nowak MA (2006) *Evolutionary Dynamics: Exploring the Equations of Life* (Belknap Press, Boston).
4. Panchanathan K, Boyd R (2003) A tale of two defectors: The importance of standing for evolution of indirect reciprocity. *J Theor Biol* 224:115–126.
5. Fehr E, Fischbacher U (2003) The nature of human altruism. *Nature* 425:785–791.
6. Fehr E, Henrich J (2003) Is strong reciprocity a maladaptation? On the evolutionary foundations of human altruism. In *Genetic and Cultural Evolution of Cooperation*, ed Hammerstein P (MIT Press, Cambridge, MA), pp 55–82.
7. Hoffman E, McCabe K, Shachat K, Smith V (1994) Preferences, property rights, and anonymity in bargaining games. *Games Econ Behav* 7:346–380.
8. Haley KJ, Fessler DMT (2005) Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evol Hum Behav* 26:245–256.

Fig. S1. The evolution of motivation (SS1; no errors possible). If repeated interactions last on average five or more rounds—surely a lower bound for human cooperative relationships that may span decades—even modest returns to cooperation generally cause agents to be very willing to cooperate when they believe that their interaction will be one-shot (SS1 simulations, using Bayesian updating). Graphs display final values of *Cooperation Probability*$_{One-Shot}$ when errors are not possible. Values are averaged over the final 500 generations within a simulation run and over distances between cue summary distributions (*d*).

Fig. S2. The evolution of motivation (SS1; errors possible). Even with the possibility of errors, if repeated interactions last on average five or more rounds, agents still evolve to be willing to cooperate with one-shot partners (SS1 simulations, using Bayesian updating). Cooperation only seems substantially reduced by the inclusion of errors when 90% of interactions are one-shot, a situation unlikely to have characterized ancestral human environments. Graphs display final values of *Cooperation Probability*_{One-Shot} when errors are possible. Values are averaged over the final 500 generations within a simulation run and over distances between cue summary distributions (*d*).

Fig. S3. The evolution of cognition (SS2; no errors possible). If repeated interactions last on average five or more rounds—likely a lower bound for human cooperative relationships—even modest returns to cooperation generally cause agents to require a great deal of evidence to believe that their interactions are one-shot (SS2 simulations, without Bayesian updating). Given these high thresholds, agents often believe their one-shot interactions are repeated and choose to cooperate. Graphs display the expected probability that agents will believe that their interactions are repeated when they are in fact one-shot. Values are averaged over the final 500 generations within a simulation run and over distances between cue summary distributions (d).

Fig. S4. The evolution of cognition (SS2; errors possible). Even with the possibility of errors, if repeated interactions last on average five or more rounds, agents still evolve to require a great deal of evidence to believe that their interactions are one-shot (SS2 simulations, without Bayesian updating). With errors present, agents generally still believe that their one-shot interactions are repeated and choose to cooperate. Cooperation only seems substantially reduced by the inclusion of errors when 90% of interactions are one-shot, a situation unlikely to have characterized ancestral human environments. Graphs display the expected probability that agents will believe that their interactions are repeated when they in fact are one-shot. Values are averaged over the final 500 generations within a simulation run and over distances between cue summary distributions (*d*).

Table S1. The evolution of motivation (SS1) regression equation (adjusted model $R^2 = 0.456$)

Predictor	b	SE	β	t	p
Intercept	−0.532	0.032		−16.370	$<10^{-57}$
Within-round benefit (b)	0.044	0.002	0.307	22.802	$<10^{-105}$
Probability of repeated interactions continuing (w)	1.315	0.031	0.577	42.847	$<10^{-300}$
Distance between cue summary distributions (d)	−0.051	0.007	−0.104	−7.693	$<10^{-13}$
Proportion of dyads that were one-shot	−0.197	0.019	−0.139	−10.323	$<10^{-23}$

Table S2. The evolution of cognition (SS2) regression equation (adjusted model $R^2 = 0.573$)

Predictor	b	SE	β	t	p
Intercept	−1.032	0.103		−9.980	$<10^{-22}$
Within-round benefit (b)	0.138	0.006	0.268	22.443	$<10^{-102}$
Probability of repeated interactions continuing (w)	3.418	0.098	0.418	35.006	$<10^{-224}$
Distance between cue summary distributions (d)	−0.124	0.021	−0.070	−5.899	$<10^{-8}$
Proportion of dyads that were one-shot	−2.890	0.061	−0.568	−47.591	$<10^{-300}$