Uncalculating cooperation is used to signal trustworthiness

Jillian J. Jordan[a,1], Moshe Hoffman[b], Martin A. Nowak[b], and David G. Rand[a,1]

[a]Department of Psychology, Yale University, New Haven, CT 06511; and [b]Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138

Humans frequently cooperate without carefully weighing the costs and benefits. As a result, people may wind up cooperating when it is not worthwhile to do so. Why risk making costly mistakes? Here, we present experimental evidence that reputation concerns provide an answer: people cooperate in an uncalculating way to signal their trustworthiness to observers. We present two economic game experiments in which uncalculating versus calculating decision-making is operationalized by either a subject's choice of whether to reveal the precise costs of cooperating (Exp. 1) or the time a subject spends considering these costs (Exp. 2). In both experiments, we find that participants are more likely to engage in uncalculating cooperation when their decision-making process is observable to others. Furthermore, we confirm that people who engage in uncalculating cooperation are perceived as, and actually are, more trustworthy than people who cooperate in a calculating way. Taken together, these data provide the first empirical evidence, to our knowledge, that uncalculating cooperation is used to signal trustworthiness, and is not merely an efficient decision-making strategy that reduces cognitive costs. Our results thus help to explain a range of puzzling behaviors, such as extreme altruism, the use of ethical principles, and romantic love.

reputation | social evaluation | decision-making | experimental economics | moral psychology

Humans are exceptional in their willingness to incur personal costs to benefit others, and a great deal of work across the social and natural sciences has sought to understand this exceptionally cooperative behavior (1–16). A central explanation that has emerged is reciprocity: often, the future benefits of cooperation outweigh the present costs, and so it is in your long-run self-interest to cooperate with others (17–21). However, there are also many contexts in which the future benefits are not sufficient to outweigh the immediate costs of cooperating: humans sometimes face the opportunity to cooperate in anonymous settings, or with strangers, or to make huge sacrifices and receive only moderate benefits in return. Thus, theories of reciprocity predict that when given the opportunity to cooperate, people should calculate the costs and benefits, and cooperate only when doing so is worthwhile. In other words, people should constantly be computing in every decision whether cooperating is worth it.

Despite this clear theoretical prediction, however, people often appear to cooperate without calculating the costs and benefits. Friends frequently grant requests to help each other without inquiring about how much time or effort will be involved, and avoid precisely tracking favors (22–24). Intimate relationships often foster strong prosocial emotions, such as devotion and love, that encourage extreme cooperative behavior that is insensitive to costs or contexts (25). People impulsively decide to help strangers in emergencies (26), and there are rich traditions of adhering to ethical principles (27) or religious teachings (28–31) that prescribe rigid guidelines for when cooperation is obligatory, regardless of the costs and benefits to the actor. These diverse examples likely evoke a broad range of proximate psychologies, ranging from intuitive and emotional processes to explicit conscious decisions not to calculate (that may themselves be the result of calculation). However, these various proximate mechanisms all lead to cooperative behavior that

is not conditional on the precise cost of cooperating in a specific situation or context—what we term "uncalculating cooperation."

When considered in the light of reciprocity, uncalculating cooperation is therefore a puzzling phenomenon that makes individuals liable to cooperate in contexts in which they would have been better off defecting. Why should people put themselves at risk of giving too much and receiving too little?

One possible explanation is that people engage in uncalculating cooperation in contexts in which they are willing to pay even the maximum possible cost of cooperation, so calculating is unnecessary. Or, relatedly, so long as cooperation is typically worthwhile, cooperating without calculating the costs can be an efficient "heuristic": it usually leads to the right decision, and avoids costs associated with calculating (e.g., cognitive costs of deliberation, or time and effort involved in gathering relevant information) (32–34), a proposal that is supported by empirical evidence (e.g., ref. 35). Here, however, we provide the first experimental evidence, to our knowledge, that uncalculating cooperation is more than just an efficient way to make cooperative decisions.

Specifically, we demonstrate that uncalculating cooperation is motivated by reputation concerns: People use uncalculating cooperation to signal their trustworthiness to observers. This hypothesis builds on evidence that people who calculate when presented with the opportunity to behave morally are perceived as less prosocial (27, 36), even when they do ultimately wind up making the "right" decision (37, 38). Calculating behavior is seen as a sign of doubt or uncertainty (37, 39, 40), whereas prosocial decisions that are quick, impulsive, or emotional are seen as reflecting genuine moral goodness (38, 41). As a result of this social cost of

Stage 1: Helping Game

Stage 2: Trust Game
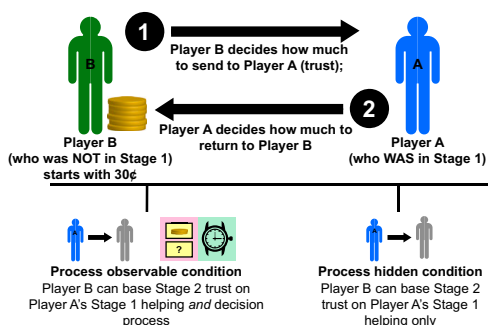
Fig. 1. Our two-stage experimental design capturing uncalculating co-operation. First, in the HG, player A has the opportunity to pay a cost to help a passive recipient. Player A decides both whether to make this decision in an uncalculated manner (operationalized via looking choice in Exp. 1 and decision time in Exp. 2) and whether to help. Second, in the TG, player B decides how much to send to player A (i.e., how much to trust), who then decides how much to return to player B (i.e., how trustworthy to be). In the process observable condition, player B can condition trust on player A's stage 1 decision process (i.e., looking choice or decision time) and helping decision. In the process hidden condition, player B can condition only on player A's stage 1 helping decision, not player A's decision process.

calculating, people may cooperate in uncalculating ways even in situations in which, absent reputational concerns, doing so would not make sense. In other words, people may engage in uncalculating cooperation, even when the maximum possible cost is not worth paying and the nonsocial costs of calculation described earlier (e.g., cognitive effort, time) are low, for the purpose of boosting their reputations. (Note that an implication of this argument is that calculations regarding the reputational costs of cooperating in a calculating manner can underlie this decision to engage in un-calculating cooperation.)

To provide empirical support for this account, we experimentally test the hypothesis that people avoid calculating the costs of co-operation because of reputational concerns. Across two experiments, we demonstrate that when people's decision-making processes are observable to others, they behave in a less calculating way. This observation suggests that they use uncalculating cooperation to gain reputational benefits, and not merely as an efficient way to avoid the (nonsocial) costs of calculating. Thus, we provide the first experimental evidence, to our knowledge, for the key prediction of the reputation account. We also show that such uncalculating co-operation pays off: Consistent with previous research (27, 36–41),

observers perceive uncalculating cooperation as a reliable signal and trust uncalculating cooperators with more money. Finally, we show that this perception is valid: People who reach the same cooperative decisions in an uncalculating way are actually more trustworthy.

In our experiments, we use a two-stage incentivized economic game (Fig. 1). Both experiments are almost identical in design, except for how they operationalize uncalculating versus calculat-ing decision-making. In the first stage of both experiments [the Helping Game (HG)], player A decides whether to pay a cost to benefit a recipient and can make this decision in a way that is calculating or uncalculating (see following for details on how "cal-culating" is operationalized in each study).

In the second stage of both experiments [the Trust Game (TG)], player B (who was not involved in the HG) receives an endowment and decides how much to send to player A. Any money sent is tripled by the experimenter. Player A then decides what percentage of this tripled amount (if any) to return to player B (without knowing how much he or she has been sent by player B). Critically, then, what is payoff-maximizing for player B depends on the be-havior of player A. It is payoff-maximizing for player B to send her entire endowment to player A if player A is trustworthy (i.e., will return a substantial amount, more than a third); conversely, if player A is untrustworthy and returns less than a third, it is payoff-maximizing for player B to send nothing. In contrast, it is never payoff-maximizing for player A to return anything to player B. Thus, the amount player B sends to player A reflects player B's trust of player A, and the amount player A returns to player B reflects player A's trustworthiness.

In the process observable condition, player B can condition trust (using the strategy method) both on player A's stage 1 decision (did player A cooperate?) and decision-making process (did player A make this decision in a calculating or uncalculating way?). In the process hidden condition, player B can condition trust only on player A's stage 1 decision, and not player A's decision-making process.

In Exp. 1 (n = 735 player A–player B pairs), we operationalize calculating behavior with a looking choice, loosely inspired by a recent formal model of "cooperating without looking" in an "en-velope game" (42, 43). (For further discussion of the relationship between our experiments and this model, see *Discussion*). In this experiment, we tell player As they have the opportunity to pay a cost to help another player, but that by default they do not know how large that cost is; they only know that it is some fraction of their starting endowment. Player As can then decide whether to help without looking at the cost (uncalculating behavior) or to look at the cost before deciding (calculating behavior). Thus, in Exp. 1, calcu-lating behavior is measured by whether or not player A chooses to look at the cost. Then, in the process observable condition, but not the process hidden condition, player B can condition stage 2 trust, via the strategy method, on whether player A looked at the cost.

In Exp. 2 (n = 737 player A–player B pairs), we operationalize calculating behavior with decision times, inspired by experimental evidence that quick cooperative choices are perceived to be more prosocial (37, 39, 40). In this experiment, we tell player As they have the opportunity to pay a cost to help another player, and that ini-tially they do not know how large that cost is, they only know that it is some fraction of their starting endowment. Then, right before they make their decisions, we reveal the exact cost of helping to player As and measure the time they spend deciding, with shorter times reflecting less calculating behavior. Thus, in Exp, 2, calculat-ing behavior is measured by the number of seconds spent making the helping decision (decision times were natural-log transformed because they were highly right-skewed, as in ref. 44). Then, in the process observable condition, but not the process hidden condition, player B can condition stage 2 trust, via the strategy method, on whether player A decided quickly (specifically, on whether player A was faster or slower than the median decision time). We note that player A deciding quickly could reflect either the use of intuitive cognition (45) or a lack of decision conflict (i.e., an "easy choice"
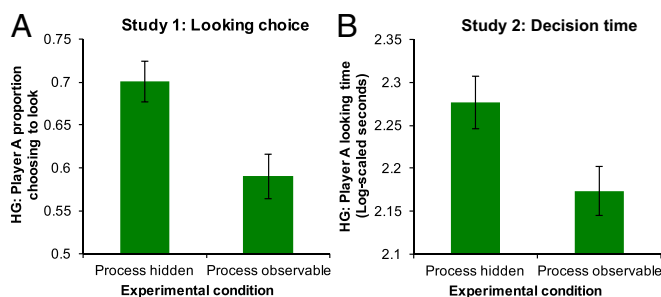
Fig. 2. Uncalculating cooperation is used as a signal of trustworthiness. Player As are more likely to engage in uncalculating behavior in the HG (stage 1) when they know their decision process will be observed by a subsequent partner in the TG (stage 2). (*A*) Results from study 1 (*n* = 735), in which we plot proportion of player As choosing to look in the HG. (*B*) Results from study 2 (*n* = 737), in which we plot natural-log transformed average decision times for player As in the HG. Error bars indicate ±1 SEM.

based on strong preference for one option over the other) (46, 47), both of which fall under the umbrella of "uncalculating" decision-making.

Within this two-stage design, our theory makes three key predictions: (*i*) people should engage in more uncalculating behavior when their decision process is observable, (*ii*) people should perceive uncalculating cooperators as more trustworthy than calculating cooperators, and (*iii*) uncalculating cooperators really should behave in a more trustworthy way than calculating cooperators.

Results

Uncalculating Cooperation Is Used as a Signal. Our key prediction is that people use uncalculating cooperation for reputational benefits: player A should be more likely to decide in an uncalculating way (and thus risk making a suboptimal decision) in the process observable condition, where doing so can confer reputational benefits, than in the process hidden condition, where doing so cannot. Indeed, we confirm this prediction in both experiments. In Exp. 1, 70% of player As looked at the cost of helping in the process hidden condition compared with only 59% in the process observable condition (*n* = 735; *P* = 0.002) (Fig. 2*A*). In Exp. 2, the mean player A decision time was 12.50 s (2.28 log-seconds) in the process hidden condition compared with 10.26 s (2.17 log-seconds) in the process observable condition (*n* = 737; *P* = 0.014) (Fig. 2*B*). Thus, we confirm our key prediction: Across two experiments, subjects behaved in a less calculating manner when their reputations were at stake. In Exp. 1, they looked less at the cost of helping when their looking choice was observable, and in Exp. 2, they decided faster when their decision time was observable.

Uncalculating Cooperation Is Perceived as a Signal. Next, we verify that uncalculating cooperation does, in fact, confer reputational benefits. In the process observable condition of both experiments, player Bs sent more to player As who reached a cooperative decision in an uncalculating way than to player As who reached a cooperative decision in a calculating way. In Exp. 1, player Bs sent an average of 55% of their endowments to player As who helped without looking at the cost compared with 49% to player As who helped after looking (*n* = 361; *P* < 0.001) (Fig. 3*A*, *Left*). In Exp. 2, player Bs sent an average of 60% of their endowments to player As who helped relatively quickly (decision time below the median) compared with 50% to player As who helped relatively slowly (decision time above the median; *n* = 365; *P* < 0.001) (Fig. 3*B*, *Left*). Thus, across two experiments, we confirmed our prediction that uncalculating cooperators are trusted more than calculating cooperators. In Exp. 1, subjects

were less trusting of individuals who checked the cost before helping, and in Exp. 2, they were less trusting of individuals who considered the cost for a long time before helping.

Furthermore, our theory predicts that uncalculating decisions should be perceived positively when they lead to cooperative behavior, specifically because they signal that the decision-maker can be trusted to cooperate in the future, and not because of a domain-general effect whereby uncalculating decisions are always desirable. Thus, we further predict that the positive effect of uncalculating decisions on trust should be specific to uncalculating cooperation and should not apply to uncalculating defection.

Indeed, in both experiments, the effect of uncalculating behavior on trust is significantly larger when the behavior in question is cooperation: When predicting player B's trust, there is a significant positive interaction between whether player A made an uncalculating decision and whether player A decided to help (Exp. 1: coefficient = 7.73; *n* = 361; *P* < 0.001; Exp. 2: coefficient = 14.74; *n* = 365; *P* < 0.001). Furthermore, uncalculating behavior is directionally negative if player A decided not to help: Player Bs trusted uncalculating defectors less than calculating defectors. In Exp. 1, player Bs sent an average of 19% of their endowments to player As who chose not to help without looking at the cost compared with 21% to player As who chose not to help after looking (*n* = 361; *P* = 0.080) (Fig. 3*A*, *Right*). In Exp. 2, player Bs sent an average of 22% of their endowments to player As who decided not to help relatively quickly (decision time below the median) compared with 27% to player As who decided not to help relatively slowly (decision time above the median; *n* = 365; *P* < 0.001) (Fig. 3*B*, *Right*). Thus, across two experiments, we confirmed our prediction that uncalculating cooperation was perceived positively, but uncalculating defection was not.

Uncalculating Cooperation Actually Is a Signal. Finally, we show that trusting uncalculating cooperators is, in fact, reasonable. Across both conditions, player As who reach a cooperative decision in an uncalculating way return more to player Bs than player As who reach a cooperative decision in a calculating way. In Exp. 1, player As who helped without looking at the cost returned an average of 50% of the amount they were sent to player B compared with 41% among player As who helped after looking (*n* = 595; *P* < 0.001) (Fig. 4*A*, *Left*). In Exp. 2, among player As who helped, there was a significant negative effect of log-transformed helping decision time on the amount returned to player B (coefficient = −4.61; *n* = 624; *P* = 0.021; because this analysis is correlational, and an individual's
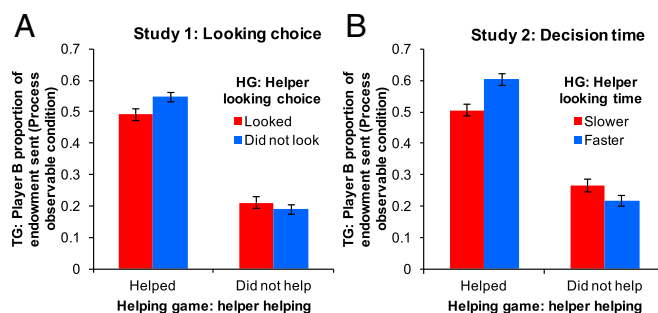
Fig. 3. Uncalculating cooperation is perceived as a signal of trustworthiness. In the process observable condition, player Bs trust player As who engaged in uncalculating cooperation more than player As who engaged in calculating cooperation, but uncalculating behavior is not perceived positively if player A did not help. (*A*) Results from study 1 (*n* = 361). (*B*) Results from study 2 (*n* = 365). In both panels, we plot average proportions of initial TG endowment sent (via the strategy method) by player B in the process observable condition, as a function of player A's decision process and helping decision in the prior stage (the HG). Error bars indicate ±1 SEM.

Jordan et al.
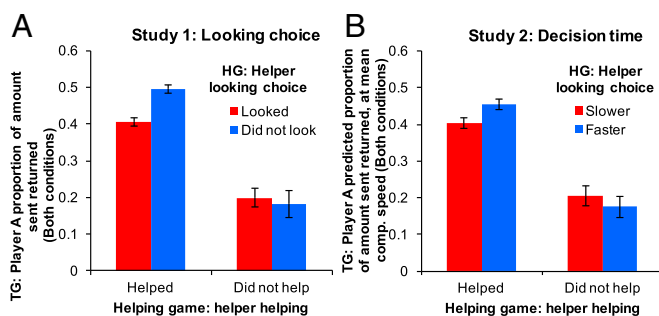
Fig. 4. Uncalculating cooperation actually is a signal of trustworthiness. Across both conditions, player As who engaged in uncalculating cooperation are more trustworthy than player As who engaged in calculating co-operation, but uncalculating behavior does not predict trustworthiness among player As who chose not to help. (*A*) Results from study 1 (*n* = 735), in which we plot average proportion returned by player A in the TG, averaged across both conditions. (*B*) Results from study 2 (*n* = 737), in which we plot predicted proportion returned by player A in the TG, based on a regression model taking data from both conditions and including natural-log trans-formed comprehension speed as a control variable. Predictions are gener-ated for a subject with a helping decision time that is either 1 SD above or below the mean. Error bars indicate ±1 SEM.

helping decision time reflects not only time spent considering the cost of helping but also general comprehension speed, this regres-sion includes a control for log-transformed time spent reading the comprehension questions; i.e., a measure of general reading/com-prehension speed) (Fig. 4*B*, *Left*). Thus, across two experiments, we confirmed our prediction that uncalculating cooperators are more trustworthy than calculating cooperators. In Exp. 1, subjects were more trustworthy if they helped without looking at the cost, and in Exp. 2, they were more trustworthy if they helped without consid-ering the cost for a long time.

Furthermore, mirroring our logic regarding player B perceptions, we expect that uncalculating decisions only predict increased trustworthiness when player A cooperates in an uncalculating way, and not when player A defects in an uncalculating way. Indeed, in both experiments, the effect of uncalculating behavior on trust-worthiness is significantly larger when the behavior in question is cooperation: when predicting player A's trustworthiness, there is a significant positive interaction between whether player A made an uncalculating decision and whether player A decided to help [Exp. 1: coefficient = 10.56; *n* = 735; *P* = 0.031; and Exp. 2 (again, controlling for general comprehension speed): coefficient = 7.69; *n* = 737; *P* = 0.019]. Furthermore, uncalculating behavior is direc-tionally (albeit nonsignificantly) negative if player A defected: Un-calculating defectors were directionally less trustworthy than calculating defectors. In Exp. 1, player As who decided not to help without looking at the cost returned an average of 18% of the amount they were sent to player B compared with 20% among player As who decided not to help after looking (*n* = 140; *P* = 0.718) (Fig. 4*A*, *Right*). In Exp. 2, among player As who decided not to help, there was a nonsignificant positive effect of log-transformed helping decision time on the amount returned to player B (co-efficient = 2.54; *n* = 113; *P* = 0.486; again, this regression includes a control for general comprehension speed) (Fig. 4*B*, *Right*). Thus, we confirmed our prediction that uncalculating decision-making only predicted trustworthiness when player A helped.

Discussion

Across two experiments, we found evidence for a reputation-based account of uncalculating cooperation: People are more likely to engage in uncalculating behavior when their decision process is ob-servable. Furthermore, we presented evidence that people perceive uncalculating cooperators (but not defectors) as more trustworthy than calculating cooperators in our paradigm, and that uncalculating

cooperators (but not defectors) really do behave in a more trust-worthy way than calculating cooperators.

Our key result, that people engage in less calculating behavior when their decision process is observable, provides the first evi-dence, to our knowledge, that people use uncalculating coopera-tion for reputational benefits, and not merely as a useful way to reduce the nonsocial costs of calculating (32–35). Although a theory of uncalculating cooperation as merely an efficient decision-making strategy can explain our second and third results (indi-viduals who cooperate across contexts to reduce the nonsocial costs of calculating will end up cooperating more, and thus should be perceived as, and should actually be, more trustworthy), it cannot explain why uncalculating decision-making should decrease when it is not observable. Based only on decision-making effi-ciency, acting in an uncalculating way should be equally valuable, regardless of who is watching. Thus, the fact that uncalculating decision-making is sensitive to observability suggests it represents a costly strategy that risks making a suboptimal choice, but has the benefit of signaling trustworthiness. This result has important im-plications for our understanding of the function of uncalculating cooperation, implicating reputational motives. It also suggests boundary conditions for when uncalculating cooperation should be observed. For example, when uncalculating cooperation serves as a reputation strategy, it should be particularly likely when trust is relatively important, as opposed to when the maximum cost of cooperation is relatively small or when the cognitive and temporal costs of calculating are relatively large.

Of course, our reputation-based account is not mutually exclusive with the idea that people are sometimes uncalculating cooperators because they are willing to pay even the maximum cost or because calculating has nonsocial costs; these accounts may help explain why some subjects cooperated in an uncalculating manner even when their decision process was hidden. However, we note that because calculating whether your decision process is observable may itself be observable to others, people may also engage in "meta uncalculat-ing cooperation" (i.e., uncalculating cooperation that is itself un-calculated, and not conditional on whether one's decision process is observable), which may also help explain uncalculating coopera-tion in the process hidden condition, and make one especially trust-worthy to others.

It is important to note that in the process observable conditions of both experiments, we explicitly inform player As that player Bs can condition their trust on looking choices/decision times, and thus that there are possible reputation consequences of calculating; conversely, in the process hidden conditions, we explain that there are no possible reputation consequences of calculating. Critically, however, this information about reputational consequences needs not be presented so explicitly to obtain our key result. In the *SI Appendix*, we present a subtler version of Exp. 1, in which we re-frain from directly telling player As what player Bs can condition their trust on (but instead convey this information indirectly via screenshots of the study from player B's perspective). This addi-tional study replicates our key finding that player As are less cal-culating when they know their decision process is observable. These results suggest that in real-world contexts of interest, when subjects are aware that their decision processes are observable (via a range of real-world observability cues, which may typically not be explicit), they are likely to act on this information by making cooperative decisions in a less calculating way. See *SI Appendix* for details.

Our second result (that people preferentially trust uncalculating cooperators) demonstrates that engaging in uncalculating decision-making can be worth the costs: Uncalculating cooperators receive reputational benefits in the form of increased trust. These findings add to a growing body of evidence showing that people attend to whether decisions are made in a calculating or uncalculating way. In one prior study, for example, subjects judged characters more posi-tively if they made prosocial decisions without hesitation, because their decisions were perceived as more certain (37), which fits with

evidence that decision time is seen as reflecting doubt across a range of social contexts (39, 40) [and evidence that decision conflict does indeed drive decision times (46, 47)]. Other studies have also found that prosocial decisions that are motivated by emotion (38) or made impulsively (41) are seen as reflecting genuine altruistic motives, that deontological decision-makers are perceived as more trustworthy (27), and individuals who decline to reveal the exact payoffs of cooperating are predicted to behave more prosocially (36). We build on this research by using incentivized economic games to demonstrate that subjects show more "revealed" trust of people who help (*i*) without looking at the cost or (*ii*) relatively quickly.

Likewise, our third result (that uncalculating cooperators really are more trustworthy) confirms that it can be beneficial to trust uncalculating cooperators: they really do return more money in the TG, suggesting that uncalculating cooperation serves as an honest signal of trustworthiness. This work builds on the finding that intuitive decisions are typically more cooperative (35) by showing that intuitive cooperation in one decision (compared with more calculated cooperation) predicts trustworthiness in a future decision.

Relatedly, our second and third results are particularly powerful, given that the target of uncalculating cooperation in the HG was a different person than the target of the trustworthiness in the TG. Uncalculating cooperation is often an important signal within dyadic relationships (e.g., a willingness to help, regardless of the costs, is a key quality of a loyal friend or romantic partner); thus, we might expect even stronger results if the recipient from the HG was the first mover in the TG. However, our results provide evidence that, to some extent, subjects expect uncalculating cooperation to predict prosociality across decisions and interaction partners (at least in a lab experiment).

Importantly, we found that uncalculating cooperation, but not uncalculating defection, is perceived as, and actually is, a positive signal of trustworthiness. Directionally, uncalculating defection is perceived as and actually is a negative signal, and these effects reach significance in some analyses ["perceived as": Exp. 2, supplemental experiment (*SI Appendix*), and pooled datasets, and marginally significant in Exp. 1; "actually is": marginally significant in supplemental experiment; see *SI Appendix* for details]. This reversal is consistent with evidence that quick decisions can be perceived positively or negatively, depending on the nature of the decision (39), and demonstrates that it is specifically uncalculating cooperation that is seen positively, rather than uncalculating decision-making generally being seen as desirable. An interesting question for future research is why uncalculating cooperation is perceived more positively (relative to calculating cooperation) than uncalculating defection is perceived negatively (relative to calculating defection). One possibility is that subjects perceived not helping as more diagnostic than helping (48), perhaps because helping could have been motivated by a desire to elicit trust, and thus were more attentive to helpers' decision processes as a result.

Another important question is how uncalculating cooperation remains an honest signal of trustworthiness. What stops people from using uncalculating cooperation to elicit trust from others, but then behaving exploitatively? One possibility comes from a model of "cooperating without looking" in the "envelope game" (42, 43). In this game, uncalculating cooperation prevents an individual from learning whether, in the current situation, defection would earn a higher payoff than cooperation. As a result, uncalculating individuals are precommitted to not knowing when defection is worthwhile, and thus will best respond to the information they have by reliably cooperating across contexts (so long as, on average, cooperation pays for them).

Another possibility is that uncalculating cooperation remains an honest indicator of trustworthiness via costly signaling (49–52). For individuals who face incentives to be trustworthy (i.e., who typically find cooperation advantageous), agreeing to cooperate without calculating is not very costly: In any given situation, it is likely that a cost–benefit analysis would support cooperation. In contrast, for individuals who face incentives to be exploitative (i.e., who rarely find cooperation advantageous), agreeing to cooperate without calculating is costlier, and it is more likely that a cost–benefit analysis would favor defection. Thus, exploitative individuals may not find uncalculating cooperation worthwhile, even when factoring in the increased trust it elicits, keeping uncalculating cooperation an honest signal.

Critically, our experiments did not explicitly build in either of these mechanisms for keeping signals honest: Nothing about our game structure could stop player As from engaging in uncalculating cooperation in the HG and then returning nothing in the TG. If we had exactly recreated the envelope game or a costly signaling model in the laboratory, purely rational subjects (without any psychological predisposition toward treating uncalculating cooperation as an honest signal of trustworthiness) would use uncalculating cooperation to signal trustworthiness and would trust uncalculating cooperators as a way to maximize their payoffs, and thus our results could merely reflect strategic reasoning in a novel game. Instead, we created a game setup in which there is not actually an "honest signaling" equilibrium, such that positive results would point to psychological predispositions regarding uncalculating cooperation (thus, for example, Exp. 1 uses a "looking choice" measure inspired by the envelope game of refs. 42 and 43, but does not formally conform to its structure). Our findings therefore suggest that human psychology has been shaped by daily life contexts in which uncalculating cooperation honestly signals trustworthiness, and that this psychology "spills over" to situations in which it does not actually make "rational" sense. Future research should further investigate the ultimate mechanisms responsible for keeping signals honest, and thus creating this psychology.

A final important future direction is investigating the proximate motivations that underlie uncalculating cooperation: How often is the choice to be uncalculating itself calculated and strategic (53)? And when uncalculated cooperation is nonstrategic, when is it deliberate (e.g., somebody consciously applying an unconditional ethical principle) versus automatic [e.g., somebody being blinded by love in an intimate relationship (25) or spontaneously helping in an emergency (26)]? Future research should investigate the proximate psychologies at play in our experiments and in daily life.

In sum, humans frequently cooperate without calculating the costs, despite the fact that doing so forfeits their ability to condition cooperation on whether or not it is worthwhile. Here, we provide the first empirical evidence, to our knowledge, that people preferentially engage in uncalculating cooperation when their reputations are at stake. Our results thus suggest that people use uncalculating cooperation to signal trustworthy to others, and not merely as an efficient decision-making strategy. These results have broad applications to behaviors ranging from extreme altruism to ethical principles to romantic love. If I help without thinking today, you know you can rely on me, no matter what tomorrow brings.

Materials and Methods

In both studies, we recruited pairs of subjects (player As and player Bs) online through Amazon Mechanical Turk to play a two-stage incentivized, anonymous economic game with no deception. We received approval from the Yale University Human Subjects Committee and gained informed consent as required by the Yale Institutional Review Board when subjects accepted our task on Amazon. In stage 1, player A participates in a HG and player B does not participate. In the HG, player A starts with 20 cents, whereas a recipient starts with nothing. Player A is informed that he or she will have a choice over whether to help the recipient by sacrificing some fraction of his or her endowment c to deliver $2c$ to the recipient. Initially, player A is not told the exact value of c, and knows only that it is between 0 and 20 cents. In Exp. 1, before making his or her helping decision, player A chooses whether or not to reveal the exact value of c (in which case she is informed that c is 10 cents). In Exp. 2, player A is always informed as to the exact value of c (10 cents), and then immediately after, makes his or her helping decision on the same screen. The time spent on this screen is recorded.

In stage 2, both player A and player B participate in a TG. In the TG, player B starts with 30 cents and chooses how much, if anything, to send to player A. Any amount sent is tripled. Player A then chooses how much, if anything, to return to player B. In the process observable condition, player B can condition his sending decision both on player A's HG decision process (whether player A revealed the cost of c in Exp. 1, and player A's decision time in Exp. 2) and on player A's HG decision (whether player A helped the recipient). In the process hidden condition, player B can only condition his sending decision on player A's HG decision. We use the strategy method for both players [i.e., player B decides how much to send to a player A, who engaged in all possible combinations of HG decision processes and/or decisions (depending on condition), without knowing what player A actually did, and player A decides what percentage of the amount player B will send to return, without knowing how much player B actually sent].

We ask subjects comprehension questions to assess their understanding of the incentive structure of both phases of the game. In our primary analyses, we report results from all subjects, but all of our results are robust to restricting to subjects who answered all comprehension questions correctly (see *SI Appendix* for details).

In our analyses, we use logistic regressions when predicting HG looking decisions (which are binary) and linear regressions when predicting HG decision times, as well as TG sending and returning decisions (which are continuous). We use robust SEs in all regressions. Player Bs make multiple TG sending decisions (because they condition their sending on different possible player A HG behaviors); we analyze these data by treating each sending decision as an observation and clustering robust SEs on subject to account for the nonindependence of repeated observations from the same subject. In our analyses of HG decision times, we natural-log transform times (because they are highly right-skewed) and control for general comprehension speed when taking decision time as an independent variable (because variance in helping decision time is likely to reflect both time spent considering the cost of helping and general compression ability). We operationalize general comprehension speed as the natural-log transformed sum of the time the subject spent on the two screens involving comprehension questions (about the two stages of our game).

1. Axelrod R, Hamilton WD (1981) The evolution of cooperation. *Science* 211(4489): 1390–1396.
2. Boyd R, Richerson PJ (1992) Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol Sociobiol* 13(3):171–195.
3. Chudek M, Henrich J (2011) Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends Cogn Sci* 15(5):218–226.
4. Crockett MJ, Kurth-Nelson Z, Siegel JZ, Dayan P, Dolan RJ (2014) Harm to others outweighs harm to self in moral decision making. *Proc Natl Acad Sci USA* 111(48): 17320–17325.
5. Galinsky A, Schweitzer M (2015) *Friend & Foe: When to Cooperate, When to Compete, and How to Succeed at Both* (Crown Business, New York).
6. Gintis H, Bowles S, Boyd R, Fehr E (2003) Explaining altruistic behavior in humans. *Evol Hum Behav* 24(3):153–172.
7. Jordan JJ, Peysakhovich A, Rand, DG (2014) Why we cooperate. *The Moral Brain: Multidisciplinary Perspectives*, eds Decety J, Wheatley T (MIT Press, Cambridge, MA), 87–101.
8. Kraft-Todd G, Yoeli E, Bhanot S, Rand D (2015) Promoting cooperation in the field. *Current Opinion in Behavioral Sciences* 3:96–101.
9. Milinski M, Semmann D, Krambeck HJ (2002) Reputation helps solve the 'tragedy of the commons'. *Nature* 415(6870):424–426.
10. Nowak MA (2006) Five rules for the evolution of cooperation. *Science* 314(5805): 1560–1563.
11. Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge Univ Press, Cambridge, UK).
12. Perc M, Szolnoki A (2010) Coevolutionary games—A mini review. *Biosystems* 99(2): 109–125.
13. Rand DG, Nowak MA (2013) Human cooperation. *Trends Cogn Sci* 17(8):413–425.
14. Van Lange PA, Otten W, De Bruin EM, Joireman JA (1997) Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *J Pers Soc Psychol* 73(4):733–746.
15. Warneken F, Tomasello M (2009) The roots of human altruism. *Br J Psychol* 100(Pt 3): 455–471.
16. Yamagishi T, et al. (2013) Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. *Organ Behav Hum Decis Process* 120(2):260–271.
17. Trivers RL (1971) The evolution of reciprocal altruism. *Q Rev Biol* 46(1):35–57.
18. Panchanathan K, Boyd R (2004) Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* 432(7016):499–502.
19. Baumard N, André J-B, Sperber D (2013) A mutualistic approach to morality: The evolution of fairness by partner choice. *Behav Brain Sci* 36(1):59–78.
20. Nowak MA, Sigmund K (2005) Evolution of indirect reciprocity. *Nature* 437(7063): 1291–1298.
21. Fudenberg D, Maskin ES (1986) The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 54(3):533–554.
22. Xue M, Silk JB (2012) The role of tracking and tolerance in relationship among friends. *Evol Hum Behav* 33(1):17–25.
23. Clark MS, Mills J (1979) Interpersonal attraction in exchange and communal relationships. *J Pers Soc Psychol* 37(1):12–24.
24. Silk JB (2003) Cooperation without counting: The puzzle of friendship. Dahlem workshop report. *Genetic and Cultural Evolution of Cooperation* (MIT Press, Cambridge, MA), pp 37–54.
25. Frank RH (1988) *Passions Within Reason: The Strategic Role of the Emotions* (WW Norton & Co, New York).
26. Rand DG, Epstein ZG (2014) Risking your life without a second thought: Intuitive decision-making and extreme altruism. *PLoS One* 9(10):e109687.
27. Everett JA, Pizarro DA, Crockett MJ (2016) Inference of trustworthiness from intuitive moral judgments. *J ExpPsychol Gen* 145(6):772–787.
28. Henrich J (2009) The evolution of costly displays, cooperation and religion: Credibility enhancing displays and their implications for cultural evolution. *Evol Hum Behav* 30(4):244–260.
29. Atkinson QD, Bourrat P (2011) Beliefs about God, the afterlife and morality support the role of supernatural policing in human cooperation. *Evol Hum Behav* 32(1):41–49.
30. Gervais WM, Norenzayan A (2012) Like a camera in the sky? Thinking about God increases public self-awareness and socially desirable responding. *J Exp Soc Psychol* 48(1):298–302.
31. Shariff AF, Norenzayan A (2007) God is watching you: Priming God concepts increases prosocial behavior in an anonymous economic game. *Psychol Sci* 18(9):803–809.
32. Bear A, Rand DG (2016) Intuition, deliberation, and the evolution of cooperation. *Proc Natl Acad Sci USA* 113(4):936–941.
33. Rand DG, et al. (2014) Social heuristics shape intuitive cooperation. *Nat Commun* 5: 3677.
34. Kiyonari T, Tanida S, Yamagishi T (2000) Social exchange and reciprocity: Confusion or a heuristic? *Evol Hum Behav* 21(6):411–427.
35. Rand DG (2016) Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychol Sci*, 10.1177/0956797616654455.
36. Capraro V, Kuilder J (2015) To know or not to know? Looking at payoffs signals selfish behavior but it does not actually mean so. Working paper (University of Amsterdam, Amsterdam). Available at papers.ssrn.com/sol3/papers.cfm?abstract_id=2679326.
37. Critcher CR, Inbar Y, Pizarro DA (2013) How quick decisions illuminate moral character. *Soc Psychol Personal Sci* 4(3):308–315.
38. Barasch A, Levine EE, Berman JZ, Small DA (2014) Selfish or selfless? On the signal value of emotion in altruistic behavior. *J Pers Soc Psychol* 107(3):393–413.
39. Van de Calseyde PP, Keren G, Zeelenberg M (2014) Decision time as information in judgment and choice. *Organ Behav Hum Decis Process* 125(2):113–122.
40. Evans AM, Van de Calseyde PPFM (June 16, 2016) The effects of observed decision time on expectations of extremity and cooperation. *J Exp Soc Psychol*, 10.1016/j.jesp.2016.05.009.
41. Pizarro D, Uhlmann E, Salovey P (2003) Asymmetry in judgments of moral blame and praise: The role of perceived metadesires. *Psychol Sci* 14(3):267–272.
42. Hoffman M, Yoeli E, Nowak MA (2015) Cooperate without looking: Why we care what people think and not just what they do. *Proc Natl Acad Sci USA* 112(6):1727–1732.
43. Hilbe C, Hoffman M, Nowak MA (2015) Cooperate without looking in a non-repeated game. *Games* 6(4):458–472.
44. Rand DG, Greene JD, Nowak MA (2012) Spontaneous giving and calculated greed. *Nature* 489(7416):427–430.
45. Kahneman D (2003) A perspective on judgment and choice: Mapping bounded rationality. *Am Psychol* 58(9):697–720.
46. Evans AM, Dillon KD, Rand DG (2015) Fast but not intuitive, slow but not reflective: Decision conflict drives reaction times in social dilemmas. *J Exp Psychol Gen* 144(5): 951–966.
47. Krajbich I, Bartling B, Hare T, Fehr E (2015) Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nat Commun* 6:7455.
48. Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD (2001) Bad is stronger than good. *Rev Gen Psychol* 5(4):323–370.
49. Zahavi A (1975) Mate selection-a selection for a handicap. *J Theor Biol* 53(1):205–214.
50. Zahavi A (1995) Altruism as a handicap: The limitations of kin selection and reciprocity. *J Avian Biol* 26(1):1–3.
51. Gintis H, Smith EA, Bowles S (2001) Costly signaling and cooperation. *J Theor Biol* 213(1):103–119.
52. Jordan JJ, Hoffman M, Bloom P, Rand DG (2016) Third-party punishment as a costly signal of trustworthiness. *Nature* 530(7591):473–476.
53. Sperber D, Baumard N (2012) Moral reputation: An evolutionary and cognitive perspective. *Mind Lang* 27(5):495–518.

Jordan et al.