Journal of Neuroscience, Psychology, and Economics
2013, Vol. 6, No. 3, 201–213

1937-321X/13/$12.00 DOI: 10.1037/npe0000010

Intelligent People Defect More in a One-Shot Prisoner's Dilemma Game

Satoshi Kanazawa and Linus Fontaine
London School of Economics and Political Science

Why so many people make the theoretically irrational decision to cooperate in a one-shot Prisoner's Dilemma game remains a puzzle in game theory. Recent developments in evolutionary psychology suggest that the anomaly may be attributable to evolutionary constraints on the human brain and their interaction with general intelligence. We conduct a laboratory experiment to test three hypotheses: (a) projection of a video image of another experimental subject increases cooperation because the human brain implicitly assumes that their choice is not anonymous; (b) more intelligent individuals are more likely to defect, because they are more likely to comprehend the evolutionarily novel features of the experiment that make defection rational; and (c) the effect of the video projection on cooperation is greater among less intelligent individuals. The experiment clearly supports two of the three hypotheses.

Keywords: evolutionary psychology, general intelligence, Savanna Principle, Savanna-IQ Interaction Hypothesis, strong reciprocity

Why so many players of one-shot Prisoner's Dilemma games choose to cooperate has been a persistent mystery in game theory. Defection strictly dominates cooperation in uniterated Prisoner's Dilemma games, with no shadow of the future or reputational effect (due to complete anonymity of choices), so cooperation is irrational. Yet roughly half the players of such one-shot Prisoner's Dilemma games choose to cooperate (Sally, 1995).

Currently, one of the most popular explanations for unilateral cooperation in social dilemma is "strong reciprocity" (Fehr & Gintis, 2007; Gintis, 2000; Gintis, Bowles, Boyd, & Fehr, 2003). The strong reciprocity theorists argue that cooperative ("other-regarding") social norms have evolved via cultural group selection, and that humans are conditionally cooperative—they are willing to cooperate, and engage in altruistic punishment of defectors at personal costs, as long as others also cooperate. These researchers contend that strong reciprocity is an adaptation.

Critics (Burnham & Johnson, 2005; Hagen & Hammerstein, 2006) point out, however, that strong reciprocity theory by itself, unsupplemented by assumptions about the evolutionary constraints of the human brain, which we outline later in the article, cannot explain cooperation in one-shot Prisoner's Dilemma and other games. Because choices in uniterated games cannot be conditional, humans, according to the strong reciprocity theory, would not cooperate in such games *if they truly understood the nature of the games.*

Among other problems, the strong reciprocity theorists cannot explain individual differences in cooperative tendency. Their own experiment (Fischbacher, Gächter, & Fehr, 2001), for example, shows that 50% of the subjects are conditional cooperators (predicted by strong reciprocity theory), whereas 30% are unconditional free riders (predicted by microeconomics). What explains such individual differences? Their own explanation is that "individual differences result from the differing ways that individuals frame a given situation, not from gen-

---

Satoshi Kanazawa and Linus Fontaine, Department of Management, London School of Economics and Political Science.

For their comments on earlier drafts, we thank Jay Belsky, Terence C. Burnham, David de Meza, Kobe Millet, Diane J. Reyniers, and Brent T. Simpson.

Correspondence concerning this article should be addressed to Satoshi Kanazawa, Department of Management, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom. E-mail: S.Kanazawa@lse.ac.uk

eralized dispositional differences" (Henrich et al., 2005, p. 814). However, this explanation simply raises another question: Why do some individuals frame the given situation in one way (to lead them to cooperate), whereas others frame it in another way (to lead them to defect)?

In this article, we present some of the recent theoretical and empirical developments in evolutionary psychology, and propose one possible explanation for why so many individuals make the theoretically irrational choice to cooperate in one-shot Prisoner's Dilemma games. The proposed explanation can further explain why some individuals are consistently and *dispositionally* more likely to cooperate across situations, whereas others are consistently and *dispositionally* more likely to defect. We then present data from one experiment that clearly support two of the three hypotheses derived from the proposed explanation.

Evolutionary Limitations and Constraints on the Human Brain

Adaptations, physical or psychological, are adapted to and designed for the conditions of the environment of evolutionary adaptedness, not necessarily to the current environment (Tooby & Cosmides, 1990). This is easiest to see in the case of physical adaptations, such as the vision and color recognition system.

What color is a banana? A banana is yellow in the sunlight and in the moonlight. It is yellow on a sunny day, on a cloudy day, on a rainy day. It is yellow at dawn and at dusk. The color of a banana appears constant to the human eye under all these conditions, despite the fact that the actual wavelengths of the light reflected by the surface of the banana under these varied conditions are different. Objectively, it is not the same color all the time. However, the human eye and color recognition system can compensate for these varied conditions, because they all occurred during the course of the evolution of the human vision system, and can perceive the objectively varied colors as constantly yellow (Cosmides & Tooby, 1999, pp. 17–19; Shepard, 1994).

So a banana looks yellow under all conditions, *except in a parking lot at night.* Under the sodium vapor lights commonly used to illuminate parking lots, a banana does not appear natural yellow. This is because the sodium vapor lights did not exist in the ancestral environment, during the course of the evolution of the human vision system, and the visual cortex is therefore incapable of compensating for them.

The same principle holds for psychological adaptations. Pioneers of evolutionary psychology (Crawford, 1993; Symons, 1990; Tooby & Cosmides, 1990) all recognize that the evolved psychological mechanisms are adapted to, and designed for, the conditions of the environment of evolutionary adaptedness, not necessarily the conditions of the current environment. Kanazawa (2004a) systematizes these observations into what he calls the Savanna Principle: *The human brain has difficulty comprehending and dealing with entities and situations that did not exist in the ancestral environment.* Burnham and Johnson (2005, pp. 130–131) refer to the same observation as *the evolutionary legacy hypothesis*, whereas Hagen and Hammerstein (2006, pp. 341–343) call it *the mismatch hypothesis*.

This essential observation can explain why some otherwise elegant scientific theories of human behavior, such as the subjective expected utility maximization theory or game theory in microeconomics, often fail empirically, because they posit entities and situations that did not exist in the ancestral environment. For example, many players of one-shot Prisoner's Dilemma games may make the theoretically irrational choice to cooperate with their partner, as we note above, possibly because the human brain has difficulty comprehending completely anonymous social exchange and absolutely no possibility of knowing future interactions (which make the game truly one-shot). Neither of these situations existed in the ancestral environment; however, they are crucial for the game-theoretic prediction of universal defection.

Fehr and Henrich (2003) suggest that one-shot encounters and exchanges might have been common in the ancestral environment. In their response to Fehr and Henrich, Hagen and Hammerstein (2006) point out that even if *one-shot* encounters might have been common in the ancestral environment, *anonymous* encounters could not have been common, and the game-theoretic prediction of defection in one-shot games requires both noniteration and anonymity. A lack of anonymity can lead to reputational concerns even in nonrepeated exchanges.

As another illustration of the Savanna Principle, individuals who watch certain types of TV shows are more satisfied with their friendships, just as they are if they had more friends or socialized with them more frequently (Derrick, Gabriel, & Hugenberg, 2009; Kanazawa, 2002). This may be because realistic images of other humans, such as TV, movies, videos, and photographs, did not exist in the ancestral environment, in which all realistic images of other humans *were* other humans. As a result, the human brain may have implicit difficulty distinguishing their "TV friends" (the characters they repeatedly see on TV shows) and their real friends.

This leads to our first hypothesis, derived from the Savanna Principle. If the human brain has difficulty truly comprehending evolutionarily novel images of other humans on video monitors, then it may implicitly interpret such images as the presence of other humans. Because the presence of other humans, who can observe one's behavior, can create reputational consequences and therefore makes cooperation potentially rational, the Savanna Principle leads us to predict that a mere presence of a video image of another human can increase cooperation in one-shot Prisoner's Dilemma games.

> *Hypothesis 1:* Subjects who make a decision in front of a video image of another human being are more likely to cooperate in a one-shot Prisoner's Dilemma game than those making the same decision in otherwise comparable condition without the video image.

Recent experiments conducted by evolutionary psychologists are consistent with Hypothesis 1. Haley and Fessler (2005) show that players are more generous in a dictator game if they make their decisions on a computer whose desktop wallpaper has a stylized picture of human eyes. Similarly, Burnham and Hare (2007) show that subjects in a public goods game contribute more if they are "watched" by a robot with human-like eyes. Numerous laboratory and field experiments have since replicated these findings (Bateson, Nettle, & Roberts, 2006; Bourrat, Baumard, & McKay, 2011; Ekström, 2012; Ernest-Jones, Nettle, & Bateson, 2011; Francey & Bergmüller, 2012; Keller & Pfattheicher, 2011; Nettle et al., 2013; Oda, Niwa, Honma, & Hiraishi, 2011; Powell, Roberts, & Nettle, 2012; Rigdon, Ishii, Watabe, & Kitayama, 2009). Note that, in all these experiments, as in ours, the actual level of anonymity of behavioral choice does not change as a result of these experimental manipulations. In all cases, the subjects' choices are completely anonymous. Yet they are more cooperative if they make their decisions in front of a picture of human eyes or a robot with human-like eyes.

The Evolution of General Intelligence

General intelligence refers to the ability to reason deductively or inductively, think abstractly, use analogies, synthesize information, and apply it to new domains (Gottfredson, 1997; Neisser et al., 1996). The concept of general intelligence poses a problem for evolutionary psychology (Chiappe & MacDonald, 2005; Cosmides & Tooby, 2002; Miller, 2000a). Evolutionary psychologists contend that the human brain consists of domain-specific evolved psychological mechanisms, which evolved to solve specific adaptive problems (problems of survival and reproduction) in specific domains. If the contents of the human brain are domain-specific, how can evolutionary psychology explain general intelligence?

In contrast to views expressed by Miller (2000b), Cosmides and Tooby (2002), and Chiappe and MacDonald (2005), Kanazawa (2004b) proposes that what is now known as general intelligence originally evolved as a domain-specific adaptation to deal with evolutionarily novel, nonrecurrent problems. The human brain consists of a large number of domain-specific evolved psychological mechanisms to solve recurrent adaptive problems. In this sense, our ancestors did not really have to *think* in order to solve such recurrent problems. Evolution has already done the thinking, so to speak, and equipped the human brain with the appropriate psychological mechanisms, which engender preferences, desires, cognitions, and emotions, and motivate adaptive behavior in the context of the ancestral environment.

Even in the extreme continuity and constancy of the ancestral environment, however, there were occasional problems that were evolutionarily novel and nonrecurrent, which required our ancestors to think and reason in order to solve. To the extent that these evolutionarily novel, nonrecurrent problems happened frequently enough in the ancestral environment (different problem each time) and had serious enough consequences for survival and reproduction, then any genetic mutation that allowed its carriers to think and reason would have been selected for, and what we now call "general intelligence" could have evolved as a domain-specific adaptation for the domain of evolutionarily novel, nonrecurrent problems. General intelligence may have become universally important in modern life (Gottfredson, 1997; Herrnstein & Murray, 1994; Jensen, 1998) only because our current environment is almost entirely evolutionarily novel. The new theory suggests, and empirical data confirm, that more intelligent individuals are better than less intelligent individuals at solving problems *only if* they are evolutionarily novel, but that more intelligent individuals are *no better* than less intelligent individuals at solving evolutionarily familiar problems, such as those in the domains of mating, parenting, interpersonal relationships, and wayfinding.

Individual Differences in the Evolutionary Constraints

The logical conjunction of the Savanna Principle and the theory of the evolution of general intelligence suggests a qualification of the Savanna Principle (Kanazawa, 2010b). If general intelligence evolved to deal with evolutionarily novel problems, then the human brain's difficulty in comprehending and dealing with entities and situations that did not exist in the ancestral environment (proposed in the Savanna Principle) should interact with general intelligence, such that the Savanna Principle holds stronger among less intelligent individuals than among more intelligent individuals. More intelligent individuals should be better able to comprehend and deal with evolutionarily novel (but *not* evolutionarily familiar) entities and situations than less intelligent individuals.

There has been accumulating evidence for this *Savanna-IQ Interaction Hypothesis* (Kanazawa, 2012). First, individuals' tendency to respond to TV characters as if they were real friends, first discovered by Kanazawa (2002), may be limited to those with below-median intelligence (Kanazawa, 2006a); individuals with above-median intelligence do not become more satisfied with their friendships by watching more TV.

Second, less intelligent individuals have more children than more intelligent individuals, even though they do not want to, possibly because they have greater difficulty effectively employing evolutionarily novel means of modern contraception (Kanazawa, 2005). Another indication that less intelligent individuals may have greater difficulty employing modern contraception effectively is the fact that the correlation between the lifetime number of sex partners and the number of children is positive among the less intelligent but negative among the more intelligent. The more sex partners less intelligent individuals have, the more children they have, as natural consequences of sexual activity; the more sex partners more intelligent individuals have, the fewer children they have, which would only be possible with the effective use of contraception.

Third, more intelligent individuals stay healthier and live longer than less intelligent individuals, possibly because they are better able to recognize and deal with evolutionarily novel threats and dangers to health in modern society (Deary, Whiteman, Starr, Whalley, & Fox, 2004; Gottfredson & Deary, 2004; Kanazawa, 2006b). For example, less intelligent children are more likely to grow up to gain weight and become obese as adults than more intelligent children are, even net of childhood social class, education, earnings, and genetic predispositions (Kanazawa, 2013).

Fourth, criminologists have long documented that, on average, criminals appear to have lower intelligence than the average population (Herrnstein & Murray, 1994; Wilson & Herrnstein, 1985). From the perspective of the Hypothesis, there are two important points to note. First, much of what we call interpersonal crime today, such as murder, assault, robbery, and theft, might have been routine means of intrasexual male competition in the ancestral environment. It may be reasonable to suggest that this might have been how men occasionally competed for resources and mating opportunities during human evolutionary history; they sometimes beat up and killed each other, and they stole from each other, even at the risk of retaliation from the victims and their allies. It may be possible to infer this from the fact that behavior that would be classified as criminal if engaged in by humans, like murder, rape, assault, and theft, are quite common among other species (Ellis, 1998), including other primates such as chimpanzees (de Waal, 1989), bonobos (de Waal, 1992), and capuchin monkeys (de Waal, Luttrell, & Canfield, 1993). However, inferring patterns of human behavior in the ancestral past from contemporary observations of nonhuman primates is always difficult and must be done very carefully and with extreme caution.

Second, the institutions that control, detect, and punish criminal behavior in society today—the police, the courts, and the prisons—are probably evolutionarily novel, for the most part; it is likely that there was very little formal third-party enforcement of norms in the ancestral environment, only second-party enforcement (victims and their kin and allies). Thus, it may make sense from the perspective of the Hypothesis that men with low intelligence may be more likely to resort to evolutionarily familiar means of competition for resources (theft rather than full-time employment) and mating opportunities (prostitution rather than computer dating), possibly because they are less likely to opt for the evolutionarily novel means, and to fully comprehend the consequences of criminal behavior imposed by evolutionarily novel entities of law enforcement.

The Savanna-IQ Interaction Hypothesis leads to two further hypotheses with respect to behavioral choice in one-shot Prisoner's Dilemma games. All the elements of a typical laboratory experiment on Prisoner's Dilemma that make defection strictly rational are evolutionarily novel. These are anonymity of choice created by the experimental conditions and guaranteed by the experimenter, and the impossibility of knowing future interaction with the partner in the experiment, which makes neither future retaliation against defection nor reputational effect possible. It is likely that neither of these conditions existed in most social exchange situations in the ancestral environment, in which such exchange was seldom truly one-shot. Virtually all instances of social exchange in the ancestral environment were likely face-to-face. The Hypothesis would therefore first predict that more intelligent subjects are more likely to comprehend these evolutionarily novel features of an experiment appropriately and defect accordingly. In contrast, less intelligent individuals are more likely to act as if one-shot Prisoner's Dilemma games are part of repeated, face-to-face social exchange typical in the ancestral environment.

> *Hypothesis 2:* More intelligent individuals are more likely to defect in one-shot Prisoner's Dilemma games than less intelligent individuals.

This prediction by the Hypothesis contrasts sharply with that by the strong reciprocity theory, which explains cooperation in social dilemmas with other-regarding preferences. Recent studies suggest that more intelligent individuals may be more likely to possess other-regarding and altruistic preferences (Deary, Batty, & Gale, 2008; Kanazawa, 2010a). Thus, the strong reciprocity theory would predict that more intelligent individuals are *more* likely to cooperate, whereas the Hypothesis would predict that they are *less* likely to cooperate.

From an evolutionary perspective, Millet and Dewitte (2007) argue that more intelligent individuals should act more altruistically than less intelligent individuals because the former are better able to recoup the resources they lose by their altruistic behavior than the latter. In other words, Millet and Dewitte suggest that altruistic behavior is a costly signal of general intelligence. Their study of a public goods game shows that "altruists," who contribute more than their share of their endowment to the public good, are significantly more intelligent than "cooperators," who contribute exactly their share, or "egoists," who contribute less than their share. However, their data show that "cooperators" are slightly (albeit nonsignificantly) *less* intelligent than "egoists."

Because Millet and Dewitte (2007) use a four-person public goods game with a stepwise production function in which participants first observe other actors nearly miss the production of public goods twice in a row (with 95% and 98% of the required contributions, respectively) before their own participation, it is not clear how directly relevant their finding is to one-shot Prisoner's Dilemma games. Further, if general intelligence evolved as a domain-specific adaptation for evolutionarily novel problems and it is not important for solving evolutionarily familiar problems, as Kanazawa (2004b) suggests, then it seems unlikely that more intelligent individuals necessarily possessed more resources in the ancestral environment.

Contrary to our Hypothesis 2, Burks, Carpenter, Goette, and Rustichini (2009) find that, as a first mover in a sequential Prisoner's Dilemma game, more intelligent individuals are more cooperative than less intelligent individuals, but the effect of intelligence becomes nonsignificant ($p > .05$) once risk preference is controlled. More intelligent individuals are less likely to be risk-averse (Benjamin, Brown, & Shapiro, in press; Frederick, 2005). Similarly, Jones (2008) shows that repeated Prisoner's Dilemma experiments conducted at universities with more intelligent students, measured by the school's average SAT and ACT scores, produce more mutual cooperation than those conducted at universities with less intelligent students. Han, Shi, Yong, and Wang (2012) show that more intelligent Chinese children do not behave more altruistically than their classmates with average or lower intelligence in public goods, ultimatum, and dictator games. Their intelligence has no effect on their behavioral choice in these games. However, more intelligent children do behave more prosocially in a complex version of the ultimatum game in which proposers and recipients swap roles.

The Savanna-IQ Interaction Hypothesis would further predict an interaction effect on behavioral choice between general intelligence and the presence of the video image. Artificial but realistic images of other human beings (such as videos, movies, and photographs) are evolutionarily novel; in the ancestral environment, all realistic images of other humans *were* other humans. The Hypothesis would therefore predict that more intelligent individuals are more likely to comprehend such evolutionarily novel stimulus as a video image and deal with it appropriately. In contrast, less intelligent individuals are more likely to respond to such an evolutionarily novel image as if it is another human being, potentially capable of finding out whether they cooperate or defect, thus creating a possible reputational effect and making cooperation potentially rational. Thus, the Hypothesis would predict that the effect of the video treatment (in Hypothesis 1) would be stronger among less intelligent individuals than among more intelligent individuals.

> *Hypothesis 3:* There is an interaction effect between the video treatment and general intelligence, such that the effect of video treatment is significantly stronger among the less intelligent individuals than among the more intelligent individuals.

We conduct an experiment to test Hypotheses 1 through 3 above.

Experiment

Participants and Procedure

Sixty-eight subjects (37 males, 31 females) participated in the experiment. They were mostly undergraduate, graduate, and summer school students at the London School of Economics and Political Science recruited individually on campus.

The subjects participated in the experiment in pairs, but did not meet each other before, during, or after the experiment. They were instructed to come to separate rooms at an appointed time, and did not interact with anyone during the experiment other than the experimenter. They were paid £5 for showing up to the experiment, and up to an additional £5 depending on the payoff from the experiment.

When they arrived, the subjects were led inside the room and given written instructions on how to play the Prisoner's Dilemma game, with the following payoff matrix (see Figure 1). Each subject was left alone in the room to read the written instructions and make their decisions. The subject was instructed to choose either "green" or "red," with the associated payoffs, and had 5 min to make the decision, during which the experimental manipulation took place (see next section). Each subject noted their de-

		Player 2	
		"Green"	"Red"
Player 1	"Green"	3, 3	0, 5
	"Red"	5, 0	1, 1

Figure 1. Payoff matrix.

cision on a piece of paper. The subject was alone in the room during the entire experiment.

After making a decision, each subject was given a test of general intelligence (see General Intelligence Test section). After the completion of the intelligence test, the subject was fully debriefed and paid according to the joint decisions made by the two players, and left the rooms at different times in order to make sure that they did not see each other.

Experimental Manipulation

The subjects were randomly assigned to two experimental conditions: control and video. In the control condition, the subject made a decision between "green" and "red" in front of a blank white screen. In the video condition, there was a projection on the white screen of an image of a 23-year-old White male, sitting in a room and playing what is clearly recognizable as the same Prisoner's Dilemma game in which the subject was participating. The man in the video faced the camera and occasionally looked up into the camera directly. It was therefore possible for the experimental subject to "meet" the eyes of the man in the video. The video projection lasted for the entire 5-min decision period. Apart from this manipulation (whether the white screen in front of which the subject makes a decision is blank or projects the image of what appears to be another subject in the experiment), the two conditions were identical in every possible way.

General Intelligence Test

Each subject, regardless of the experimental condition, took a Raven's-type nonverbal test of general intelligence, obtained from the Web site of the German Mensa (http://www.mensa.de/index.php?id=65). The subject had 20 min to answer 33 questions. The number of correct answers was recorded for each subject.

Unfortunately, the German Mensa does not provide the norming table for its test. We therefore use the norming data for college graduates ages 20 to 30 from Wonderlic (2002). They assume the mean of 120 and standard deviation of 13.27 for this population, and we adopted their norming data for our sample. Because norming is a purely linear transformation (in this particular instance, $IQ = 120 + 13.27^{*}Z_{score}$), the particular norming data that we used did not at all affect our substantive conclusions about the effect of general intelligence on behavioral choice in Prisoner's Dilemma games. They only affected the computation of mean IQ by behavioral choice.

Results

Hypothesis 1

Table 1 (Column 1) presents the results of a logistic regression analysis in which the subject's binary behavioral choice (0 = defection, 1 = cooperation) was regressed on subject's sex (0 = female, 1 = male), treatment condition (0 = control, 1 = video), and subject's IQ. Controlling for treatment condition and general intelligence, sex had no effect on the behavioral choice ($b = -.1334$, *ns*, $e^b = .8751$); men and women were equally likely to cooperate in the Prisoner's Dilemma game. In sharp contrast, controlling for sex and general intelligence, the treatment condition had a very large and statistically significant effect on cooperation ($b = 1.0571$, $p < .05$, $e^b = 2.8780$). Subjects who made their decision in front of a video image of what appears to be another subject in the same experiment had *nearly three times the odds* of cooperation as those in the control condition, who made their decision in front of a blank screen. This result supports Hypothesis 1, derived from the Savanna Principle. The subjects appeared to act as if the video image of another human being in front of them *was* another human being.

Figure 2 presents the mean cooperation rates of subjects in the two treatment conditions. It shows that those in the control condition had the mean cooperation rate of .35, whereas those in the experimental (video) condition had the mean cooperation rate of .56. In other words, two-thirds of the subjects in the control condition made the theoretically rational decision to defect in a one-shot Prisoner's Dilemma game, whereas a majority of those in the experimental condition made the theoretically irrational decision to cooperate.

Hypothesis 2

Table 1 (Column 1) also shows that the subject's general intelligence had a significantly negative effect on cooperation ($b = -.0438$,

Table 1
The Effects of Video and IQ on Cooperation (0 = Defection, 1 = Cooperation)

	(1) Full sample	(2) Below-median intelligence	(3) Above-median intelligence
Sex	−.1334	−1.1464	.1079
(0 = female, 1 = male)	(.5186)	(.8735)	(.8083)
	.8751	*.3178*	*1.1139*
Treatment	1.0571*	2.3660*	.0155
(0 = control, 1 = video)	(.5312)	(1.0022)	(.8465)
	2.8780	*10.6542*	*1.0156*
IQ	−.0438*	−.2323**	.0412
	(.0209)	(.0865)	(.0526)
	.9571	*.7927*	*1.0420*
Constant	4.6117	25.4437	−6.0542
	(2.4700)	(9.5942)	(6.7092)
Cox & Snell pseudo-R^2	.1099	.3658	.0285
−2 log likelihood	85.8225	35.2708	38.5630
n	68	38	30

Note. Main entries are unstandardized regression coefficients. Numbers in parentheses are standard errors. Italicized entries are partial effects on odds (e^b).
* $p < .05$. ** $p < .01$.

$p < .05$, $e^b = .9571$). The odds ratio of .9571 means that an increase of 1 point in IQ decreases the odds of cooperation by more than 4%. A one-standard-deviation increase in IQ *nearly halves* the odds of cooperation ($e^{(-.0438*15)} = .5184$). The significantly negative effect of general intelligence on cooperation supports Hypothesis 2, derived from the Savanna-IQ Interaction Hypothesis, and contradicts the prediction by the strong reciprocity theory. Despite having more altruistic and other-regarding preferences, more intelligent individuals are more likely to defect in a one-shot Prisoner's Dilemma game than less intelligent individuals.

Figure 3 presents the mean cooperation rates among subjects who are below median in general intelligence ($n = 38$) and those who are above the median ($n = 30$). (There are nine subjects in the median category included in the below-median category, hence the uneven split. All of our substantive conclusions remain unchanged if we categorize these nine subjects in the above-median group.) It shows that a ma-

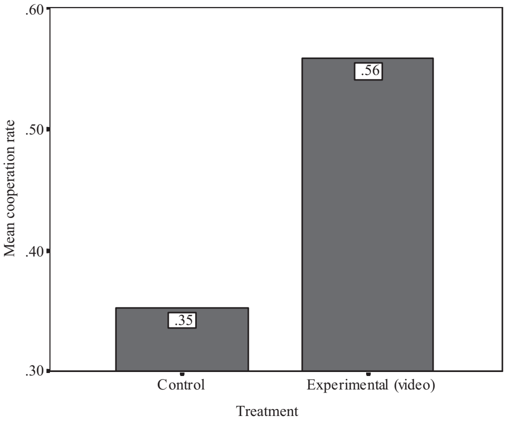

Figure 2. Mean cooperation × experimental condition.

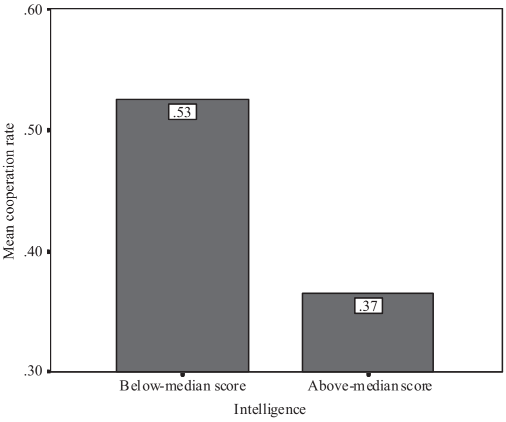

Figure 3. Mean cooperation × general intelligence.

jority (53%) of subjects below median in general intelligence made the theoretically irrational choice to cooperate in a one-shot Prisoner's Dilemma game, whereas nearly two-thirds (63%) of those above median in general intelligence made the theoretically rational decision to defect.

Figure 4 presents the data in an alternative manner. It compares the mean IQ of subjects who chose to defect and that of those who chose to cooperate. It shows that, using the norming data from Wonderlic (2002), the mean IQ of defectors is 122.8, whereas that of cooperators is 116.7.

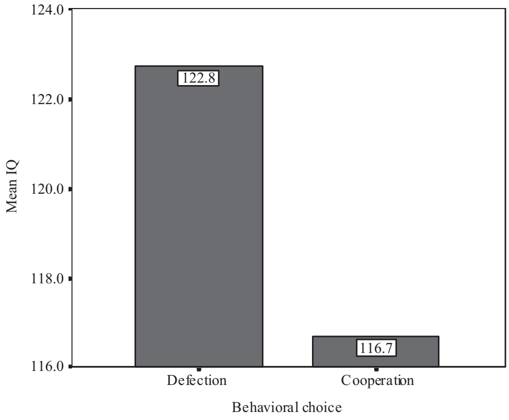

Figure 4. Mean IQ × behavioral choice.

Hypothesis 3

The most straightforward way to test Hypothesis 3—the prediction that more intelligent subjects respond less to the video manipulation in the experiment than less intelligent subjects—is to include an interaction term between intelligence and treatment as a predictor in the binary logistic regression equation. However, the treatment variable and the interaction term are highly collinear ($r = .9865$), and thus both cannot be included in the equation simultaneously.

As an alternative method of testing Hypothesis 3, we first divide the subjects into two groups: those whose intelligence is below the median ($n = 38$) and those whose intelligence is above the median ($n = 30$). We then estimate the same binary logistic regression separately for the two groups, to see if the effects of experimental treatment differ by intelligence.

As Table 1 (Column 2) shows, among the subjects with below-median intelligence, the experimental treatment has a large and statistically significant effect on cooperation ($b = 2.3660$, $p < .05$, $e^b = 10.6542$); among this group, the exposure to the video increased the odds of cooperation *by nearly 11 times* (although we must be careful in extrapolating the result due to the small sample size). Among the subjects with above-median intelligence, the experimental treatment does not have a significant effect on cooperation ($b = .0155$, *ns*, $e^b = 1.0156$). The absolute difference in the magnitude of the two coefficients is very large (2.3660 vs. .0155).

Figure 5 graphically depicts the interaction effect of the video treatment and general intelligence on cooperation. It shows that the difference in cooperation rates between the experimental and control groups is much larger among subjects below median in general intelligence than among subjects above median in general intelligence. The mean cooperation rate among subjects with below-median general intelligence in the experimental (video) condition is .68; nearly 70% of these subjects made the theoretically irrational choice to cooperate. In contrast, the mean cooperation rate among subjects with above-median general intelligence in the control condition is .33; nearly 70% of these subjects made the theoretically rational choice to defect.

However, due possibly to very small sample size and, consequently, large standard errors for the coefficients, the 95% confidence intervals for the two coefficients overlap (below-median: $.4017 < b < 4.3302$; above-median: $-1.6437 < b < 1.6747$). The large absolute difference in the slopes between the two IQ groups is therefore *not* statistically significant at .05. So we must conclude that there is no statistical evidence to support Hypothesis 3.

Discussion

In their devastating critique of the expanding literature on "strong reciprocity" by a group of strong reciprocity theorists whom they dub "the Collective," Burnham and Johnson (2005, p. 131) state

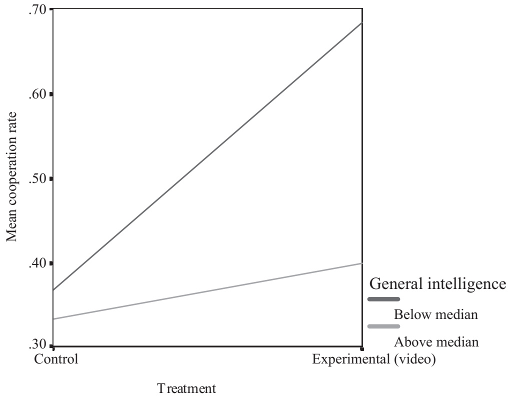

Figure 5. The interaction effect between general intelligence and video on behavioral choice.

> any experiment that (a) removes all conscious *expectation* of future rewards, reputation or kin, *and* (b) finds that SR [strong reciprocity] nevertheless varies with the subconscious *perception* of cues for these same three factors is, first, a direct falsification of the Collective's view and, second, support for our own [the evolutionary legacy hypothesis].

We believe that our experiment satisfies both conditions specified by Burnham and Johnson, and thus provides empirical support for their evolutionary legacy hypothesis and our Savanna Principle, which are substantively identical.

In our experiment, the subjects *consciously* knew that (a) the game is strictly one-shot and is never repeated (and there is therefore no possibility of "future rewards" in the form of reciprocal altruism and thus no "shadow of the future"); (b) they would never meet their exchange partner after the experiment, and their behavioral choice was completely anonymous, so there are no reputational effects; and (c) their exchange partner was a fellow student and was therefore not kin. Nevertheless, we find that the subject's behavioral choice (cooperation vs. defection) varied significantly as a function of *subconscious perception* of cues to possible reputational effect (in the form of a video image of another subject in the experiment). It seems difficult to explain the very strong effect of the experimental manipulation on cooperation in purely economic (rational choice) terms, without invoking some sort of evolutionary constraints on the human brain posited by the Savanna Principle (Kanazawa, 2004a), the evolutionary legacy hypothesis (Burnham & Johnson, 2005), or the mismatch hypothesis (Hagen & Hammerstein, 2006). The strong effect of our experimental manipulation is consistent with a large number of laboratory and field experiments conducted by evolutionary psychologists (Bateson et al., 2006; Bourrat et al., 2011; Burnham & Hare, 2007; Ekström, 2012; Ernest-Jones et al., 2011; Francey & Bergmüller, 2012; Haley & Fessler, 2005; Keller & Pfattheicher, 2011; Nettle et al., 2013; Oda et al., 2011; Powell et al., 2012; Rigdon et al., 2009). These experiments, together with ours, provide strong support for the Savanna Principle.

In addition, consistent with the Savanna-IQ Interaction Hypothesis, our results show that more intelligent individuals are significantly less likely to cooperate in a one-shot Prisoner's Dilemma game. This is possibly because individuals with higher general intelligence are better able to comprehend evolutionarily novel features of the experiment, such as complete anonymity of their choice, and impossibility of knowing future interaction with their exchange partner in the experiment, whereas individuals with lower general intelligence are less able to comprehend such evolutionarily novel situations. Note that without the evolutionarily novel features of the experiment, cooperation *is* potentially rational, as it would have been in most situations of social exchange in the ancestral environment. The Hypothesis suggests, and our results confirm, that less intelligent individuals are significantly more likely to cooperate in a one-shot Prisoner's Dilemma game than more intelligent individuals. Further consistent with the Hypothesis, the effect of the experimental manipulation of a video image of another player tends to be stronger among less intelligent subjects than among more intelligent subjects. However, this interaction effect, although substantively very large, is not statistically significant, possibly due to the small sample size.

One possible objection to our conclusion is that, being students of a highly selective elite university, all of our subjects, even in the below-median group, were highly intelligent individuals. For example, using the Wonderlic (2002) norming data, only three of our 68 subjects had IQs below 100. However, we note that virtually all of the past experiments on one-shot Prisoner's Dilemma and other games have used university students, and the conclusion that only about half of the subjects make the theoretically rational choice to defect comes from these experiments (Sally, 1995). We propose (and our data presented in Figure 5 suggest) that even larger proportions of players than the typical half may choose to cooperate in one-shot Prisoner's Dilemma games if they represent the entire range of the IQ distribution.

References

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters, 2,* 412–414. doi:10.1098/rsbl.2006.0509

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (in press). Who is "behavioral"? Cognitive ability and anomalous preferences. *Journal of the European Economic Association*.

Bourrat, P., Baumard, N., & McKay, R. (2011). Surveillance cues enhance moral condemnation. *Evolutionary Psychology, 9,* 193–199.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *PNAS, Proceedings of the National Academy of Sciences of the United States of America, 106,* 7745–7750. doi:10.1073/pnas.0812360106

Burnham, T. C., & Hare, B. (2007). Engineering human cooperation: Does involuntary neural activation increase public goods contribution? *Human Nature, 18,* 88–108. doi:10.1007/s12110-007-9012-2

Burnham, T. C., & Johnson, D. D. P. (2005). The biological and evolutionary logic of human cooperation. *Analyse & Kritik, 27,* 113–135.

Chiappe, D., & MacDonald, K. (2005). The evolution of domain-general mechanisms in intelligence and learning. *Journal of General Psychology, 132,* 5–40. doi:10.3200/GENP.132.1.5-40

Cosmides, L., & Tooby, J. (1999). *What is evolutionary psychology?* Unpublished manuscript, Center for Evolutionary Psychology, University of California, Santa Barbara.

Cosmides, L., & Tooby, J. (2002). Unraveling the enigma of human intelligence: Evolutionary psychology and the multimodular mind. In R. J. Sternberg & J. C. Kaufman (Eds.), *The evolution of intelligence* (pp. 145–198). Mahwah, NJ: Erlbaum.

Crawford, C. B. (1993). The future of sociobiology: Counting babies or proximate mechanisms? *Trends in Ecology and Evolution, 8,* 183–186. doi:10.1016/0169-5347(93)90145-F

Deary, I. J., Batty, G. D., & Gale, C. R. (2008). Bright children become enlightened adults. *Psychological Science, 19,* 1–6. doi:10.1111/j.1467-9280.2008.02036.x

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology, 86,* 130–147. doi:10.1037/0022-3514.86.1.130

Derrick, J. L., Gabriel, S., & Hugenberg, K. (2009). Social surrogacy: How favored television programs provide the experience of belonging. *Journal of Experimental Social Psychology, 45,* 352–362. doi:10.1016/j.jesp.2008.12.003

Ekström, M. (2012). Do watching eyes affect charitable giving? Evidence from a field experiment. *Experimental Economics, 15,* 530–546. doi:10.1007/s10683-011-9312-6

Ellis, L. (1998). Neodarwinian theories of violent criminality and antisocial behavior: Photographic evidence from nonhuman animals and a review of the literature. *Aggression and Violent Behavior, 3,* 61–110. doi:10.1016/S1359-1789(96)00019-5

Ernest-Jones, M., Nettle, D., & Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: A field experiment. *Evolution and Human Behavior, 32,* 172–178, doi:10.1016/j.evolhumbehav.2010.10.006

Fehr, E., & Gintis, H. (2007). Human motivation and social cooperation: Experimental and analytical foundations. *Annual Review of Sociology, 33,* 43–64. doi:10.1146/annurev.soc.33.040406.131812

Fehr, E., & Henrich, J. (2003). Is strong reciprocity a maladaptation? On the evolutionary foundations of human altruism. In P. Hammerstein (Ed.), *Genetic and cultural evolution of cooperation* (pp. 55–82). Cambridge, MA: MIT Press. doi:10.2139/ssrn.382950

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters, 71,* 397–404. doi:10.1016/S0165-1765(01)00394-9

Francey, D., & Bergmüller, R. (2012). Images of eyes enhance investments in a real-life public good. *PLoS One, 7,* e37397. doi:10.1371/journal.pone.0037397

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19,* 25–42. doi:10.1257/089533005775196732

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology, 206,* 169–179. doi:10.1006/jtbi.2000.2111

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior, 24,* 153–172. doi:10.1016/S1090-5138(02)00157-5

Gottfredson, L. S. (1997). Why *g* matters: The complexity of everyday life. *Intelligence, 24,* 79–132. doi:10.1016/S0160-2896(97)90014-3

Gottfredson, L. S., & Deary, I. J. (2004). Intelligence predicts health and longevity, but why? *Current Directions in Psychological Science, 13,* 1–4. doi:10.1111/j.0963-7214.2004.01301001.x

Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical Population Biology, 69,* 339–348. doi:10.1016/j.tpb.2005.09.005

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26,* 245–256. doi:10.1016/j.evolhumbehav.2005.01.002

Han, R., Shi, J., Yong, W., & Wang, W. (2012). Intelligence and prosocial behavior: Do smart children really act nice? *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues, 31,* 88–101. doi:10.1007/s12144-012-9133-6

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., . . . Tracer, D. (2005). ”Economic man” in cross-cultural perspective: Behavioral experiments in 15 small scale societies. *Behavioral and Brain Sciences, 28,* 795–815. doi:10.1017/S0140525X05000142

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Jones, G. (2008). Are smarter groups more cooperative? Evidence from prisoner's dilemma experiments, 1959–2003. *Journal of Economic Behavior & Organization, 68,* 489–497. doi:10.1016/j.jebo.2008.06.010

Kanazawa, S. (2002). Bowling with our imaginary friends. *Evolution and Human Behavior, 23,* 167–171. doi:10.1016/S1090-5138(01)00098-8

Kanazawa, S. (2004a). The Savanna Principle. *Managerial and Decision Economics, 25,* 41–54. doi:10.1002/mde.1130

Kanazawa, S. (2004b). General intelligence as a domain-specific adaptation. *Psychological Review, 111,* 512–523. doi:10.1037/0033-295X.111.2.512

Kanazawa, S. (2005). An empirical test of a possible solution to "the central theoretical problem of human sociobiology." *Journal of Cultural and Evolutionary Psychology, 3,* 255–256. doi:10.1556/JCEP.3.2005.3-4.3

Kanazawa, S. (2006a). Why the less intelligent may enjoy television more than the more intelligent. *Journal of Cultural and Evolutionary Psychology, 4,* 27–36. doi:10.1556/JCEP.4.2006.1.2

Kanazawa, S. (2006b). Mind the gap . . . in intelligence: Reexamining the relationship between inequality and health. *British Journal of Health Psychology, 11,* 623–642. doi:10.1348/135910705X69842

Kanazawa, S. (2010a). Why liberals and atheists are more intelligent. *Social Psychology Quarterly, 73,* 33–57.

Kanazawa, S. (2010b). Evolutionary psychology and intelligence research. *American Psychologist, 65,* 279–289. doi:10.1037/a0019378

Kanazawa, S. (2012). *The intelligence paradox: Why the intelligent choice isn't always the smart one*. New York, NY: Wiley.

Kanazawa, S. (2013). Childhood intelligence and adult obesity. *Obesity, 21,* 434–440. doi:10.1002/oby.20018

Keller, J., & Pfattheicher, S. (2011). Vigilant self-regulation, cues of being watched and cooperativeness. *European Journal of Personality, 25,* 363–372. doi:10.1002/per.797

Miller, G. F. (2000a). How to keep our metatheories adaptive: Beyond Cosmides, Tooby, and Lakatos. *Psychological Inquiry, 11,* 42–46.

Miller, G. F. (2000b). Sexual selection for indicators of intelligence. In G. R. Bock, J. A. Goode, & K. Webb (Eds.), *The nature of intelligence* (pp. 260–275). New York, NY: Wiley. doi:10.1002/0470870850.ch16

Millet, K., & Dewitte, S. (2007). Altruistic behavior as a costly signal of general intelligence. *Journal of Research in Personality, 41,* 316–326.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., . . . Urbina, S. (1996). Intelligence: Known and unknowns. *American Psychologist, 51,* 77–101. doi:10.1037/0003-066X.51.2.77

Nettle, D., Harper, Z., Kidson, A., Stone, R., Penton-Voak, I. S., & Bateson, M. (2013). The watching eyes effect in the Dictator Game: It's not how much you give, it's being seen to give something. *Evolution and Human Behavior, 34,* 35–40. doi:10.1016/j.evolhumbehav.2012.08.004

Oda, R., Niwa, Y., Honma, A., & Hiraishi, K. (2011). An eye-like painting enhances the expectation of a good reputation. *Evolution and Human Behavior, 32,* 166–171. doi:10.1016/j.evolhumbehav.2010.11.002

Powell, K. L., Roberts, G., & Nettle, D. (2012). Eye images increase charitable donations: Evidence from an opportunistic field experiment in a supermarket. *Ethology, 118,* 1096–1101. doi:10.1111/eth.12011

Rigdon, M., Ishii, K., Watabe, M., & Kitayama, S. (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology, 30,* 358–367. doi:10.1016/j.joep.2009.02.002

Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and Society, 7,* 58–92. doi:10.1177/1043463195007001004

Shepard, R. N. (1994). Perceptual-cognitive universals as reflections of the world. *Psychonomic Bulletin & Review, 1,* 2–28. doi:10.3758/BF03200759

Symons, D. (1990). Adaptiveness and adaptation. *Ethology & Sociobiology, 11,* 427–444. doi:10.1016/0162-3095(90)90019-3

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology & Sociobiology, 11,* 375–424. doi:10.1016/0162-3095(90)90017-Z

de Waal, F. B. M. (1989). Food sharing and reciprocal obligations among chimpanzees. *Journal of Human Evolution, 18,* 433–459. doi:10.1016/0047-2484(89)90074-2

de Waal, F. B. M. (1992). Appeasement, celebration, and food sharing in the two *Pan* species. In T. Nishida, W. C. McGrew, & P. Marler (Eds.), *Topics in primatology: Human origins* (pp. 37–50). Tokyo, Japan: University of Tokyo Press.

de Waal, F. B. M., Luttrell, L. M., & Canfield, M. E. (1993). Preliminary data on voluntary food sharing in brown capuchin monkeys. *American Journal of Primatology, 29,* 73–78. doi:10.1002/ajp.1350290108

Wilson, J. Q., & Herrnstein, R. J. (1985). *Crime and human nature: The definitive study of the causes of crime*. New York, NY: Touchstone.

Wonderlic. 2002. *Wonderlic personnel test & scholastic level exam user's manual*. Libertyville, IL: Author.

Received March 2, 2013
Revision received May 25, 2013
Accepted June 3, 2013 ■