



## Short Communication

## Extraversion, neuroticism, and the prisoner's dilemma

Jacob B. Hirsh\*, Jordan B. Peterson

Department of Psychology, University of Toronto, Sidney Smith Hall, 100 St. George St., Toronto, Ontario, Canada M5S 3G3

## ARTICLE INFO

## Article history:

Received 4 August 2008

Received in revised form 14 September 2008

Accepted 7 October 2008

## Keywords:

Personality

Big Five

Prisoner's dilemma

Experimental economics

Decision-making

Economic psychology

## ABSTRACT

The prisoner's dilemma has been used to study self-interest and cooperation in a variety of contexts. Applying an individual differences approach to this paradigm allows for the examination of dispositional factors that predict the likelihood of betraying one's game partner. An iterative prisoner's dilemma was administered to undergraduate students, along with measures of demographics, personality, and cognitive ability. Results demonstrate that higher scores on the withdrawal aspect of neuroticism and the enthusiasm aspect of extraversion independently predicted a greater likelihood of cooperation.

© 2008 Elsevier Ltd. All rights reserved.

## 1. Introduction

Experimental economics presents a framework for studying human decision-making, allowing theoretical models of economic behavior to be tested directly (Davis & Holt, 1993). One widely-employed paradigm in this field is the prisoner's dilemma (Rapoport & Chammah, 1965). Although the framing of this game varies substantially, the basic structure is as follows: two individuals are faced with a choice between cooperation and defection. If both parties defect, then both receive a negative outcome. If only one party defects, that party receives a very positive outcome, while the other receives a very negative outcome. Finally, if both parties cooperate, they both receive a moderately positive outcome. Although players can always gain better outcomes for themselves in the short-term by defecting, people tend to have mixed decision profiles (Cooper, DeJong, Forsythe, & Ross, 1996). Cooperation in the prisoner's dilemma appears to emerge after a number of iterative trials with the same players (Axelrod & Hamilton, 1981; Killingback, Doebeli, & Knowlton, 1999; Kreps, Milgrom, Roberts, & Wilson, 1982). This cooperation even occurs when participants are randomly matched with anonymous partners (Ellison, 1994). Cooperation tends to produce better outcomes for both players (Bendor, Kramer, & Stout, 1991), although the best strategies involve flexible response patterns (Nowak & Sigmund, 1993).

Approaching this paradigm from an individual differences approach may reveal the dispositional characteristics associated with

different response patterns. Previous research suggests that personality can indeed play an important role in the prisoner's dilemma (Boone, De Brabander, & van Witteloostuijn, 1999; Kuhlman & Marshello, 1975), although this research did not employ the five factor model, which has become the standard trait framework for research in personality (Goldberg, 1993). In the current study, we examine how decisions in the prisoner's dilemma are related to cognitive ability, demographic variables, and the Big Five personality traits of extraversion, agreeableness, conscientiousness, neuroticism, and openness/intellect. An additional goal was to examine whether such decisions correlate with mid-level personality traits, which provide greater fidelity than the broad Big Five dimensions.

## 2. Methods

## 2.1. Participants

Participants included 77 undergraduate students from the University of Toronto (60 female), with an age range of 17–39 years ( $M = 19.4$ ,  $SD = 2.8$ ). Students were recruited from introductory Psychology classes and were given course credit for their participation. The sample consisted mostly of students from European Canadian (53.2%) and East Asian (26.0%) backgrounds.

## 2.2. Materials

## 2.2.1. Prisoner's dilemma

A computerized version of the prisoner's dilemma was administered using the z-Tree software package (Fischbacher, 2007). Each

\* Corresponding author. Tel.: +1 416 946 5812; fax: +1 416 978 4811.  
E-mail address: [jacob.hirsh@utoronto.ca](mailto:jacob.hirsh@utoronto.ca) (J.B. Hirsh).

trial involved two competing players. At the beginning of each round, players were given the classic prisoner's dilemma scenario, in which both they and their partners in crime were being interrogated in separate rooms. The participants could choose to blame their partners or to say nothing. If both participants blamed the other, harsh fines were imposed. If only one person blamed the other, he would receive a small fine while the partner received a large fine. Finally, if both participants blamed each other, both received a large fine. Each game lasted 10 sequential trials.

### 2.2.2. Big Five aspect scale (BFAS, DeYoung, Quilty, & Peterson, 2007)

The BFAS is an empirically-validated instrument for measuring the broad Big Five dimensions of personality, as well as their constituent aspects. The aspects of the Big Five reflect mid-level personality traits, situated between the five broad domains and the narrower facet level of personality. Each Big Five domain breaks down into two of these mid-level aspects, including assertiveness and enthusiasm (extraversion), compassion and politeness (agreeableness), industriousness and orderliness (conscientiousness), volatility and withdrawal (neuroticism), and openness and intellect (openness/intellect). Although the aspects from each domain are correlated with each other, they also show divergent validity. The BFAS thus provides a good assessment of the broad Big Five domains, and provides the additional advantage of assessing empirically-derived mid-level personality traits.

The questionnaire features 100 descriptions with which respondents rate their agreement on a 5-point Likert scale. All items were validated as markers for the 10 mid-level Big Five aspects. As a measure of the broad domains, the BFAS has been validated against standard Big Five instruments such as the BFI (John & Srivastava, 1999) and the NEO PI-R (Costa & McCrae, 1992) with an average uncorrected correlation of  $r = .76$ . The scale also demonstrates internal (mean  $r = .83$ ) and test-retest (mean  $r = .81$ ) reliability.

### 2.2.3. Wonderlic® personnel test (WPT, Wonderlic, 1983)

The WPT was used as a measure of cognitive ability. The 50 items on the WPT are based on those from the original Otis Test of Mental Ability, and results correlate highly with full-scale IQ as measured by the Wechsler Adult Intelligence Scale, Third Edition ( $r = .92$ ; Dodrill, 1981; Dodrill & Warner, 1988; McKelvie, 1989).

## 2.3. Procedure

Participants signed a consent form upon entering the lab. They were then brought into separate rooms for computerized testing. Each participant completed the BFAS, WPT, and a demographics form. Participants were told that they would play an economic game against another randomly assigned participant, who was seated in a different room. All participants were told that they could not be identified by the other player. It was also explained that participants could earn extra money for performing well. During each session, two participants would play the game simultaneously. If only one participant was available, a confederate took

the role of the second player and was instructed to counterbalance strategies across participants. The parameters of the prisoner's dilemma were explained to the participants, who were then left to play 10 sequential trials of the game. At the end of the experiment, participants were debriefed and given \$5 in addition to course credit.

## 3. Results

Descriptive statistics are presented in Table 1. Of the 770 trials, 49% of the decisions were to cooperate and 51% were to defect. Confederates acted as the second player in 29 of the sessions, and their data was excluded from all analyses. Playing against a confederate had no significant influence on participants' defection rates,  $t(768) = 1.17$ ,  $p > .05$ . We conducted a hierarchical multiple regression to examine the factors that predicted whether participants would defect on any given trial. In the first regression block, we entered socioeconomic status (SES), the number of years that the participant had been speaking English, and whether or not the other player defected on the previous trial. In the second regression block, a stepwise analysis was conducted with cognitive ability and the broad Big Five personality domains. Participants were significantly more likely to defect if their partner defected on the previous trial ( $\beta = .27$ ,  $t(689) = 7.31$ ,  $p < .001$ ). Defection occurred on 41% of trials following partner cooperation, and 66% of trials following partner defection. Individuals with higher SES were also less likely to defect ( $\beta = -.14$ ,  $t(689) = -3.47$ ,  $p < .001$ ), but no effects were observed for years English ( $\beta = .04$ ,  $t(689) = 1.06$ ,  $p = .29$ ). In terms of dispositional variables, only neuroticism emerged as a significant predictor, with more neurotic individuals being less likely to defect ( $\beta = -.12$ ,  $t(689) = -3.08$ ,  $p < .005$ ). The overall model accounted for a significant amount of the trial-by-trial variance in decision-making,  $R^2 = .09$ ,  $F(4, 689) = 15.92$ ,  $p < .001$ .

In order to examine the role of mid-level personality traits, we repeated the hierarchical regression described above, but replaced the broad Big Five dimensions with the 10 Big Five aspects in the second-block stepwise regression. Similar effects were observed for SES ( $\beta = -.12$ ,  $t(688) = -3.06$ ,  $p < .005$ ), years English ( $\beta = .08$ ,  $t(688) = 1.90$ ,  $p = .06$ ), and the opponent's previous decision ( $\beta = .27$ ,  $t(688) = 7.40$ ,  $p < .001$ ). Personality effects were observed for the withdrawal aspect of neuroticism ( $\beta = -.14$ ,  $t(688) = -3.66$ ,  $p > .001$ ), and the enthusiasm aspect of extraversion ( $\beta = -.12$ ,  $t(688) = -2.96$ ,  $p > .005$ ), such that higher levels of these traits predicted greater cooperation. The overall model was significant, accounting for slightly more variance in decision-making,  $R^2 = .10$ ,  $F(4, 688) = 14.67$ ,  $p < .001$ .

## 4. Discussion

Individual difference variables were found to predict the likelihood of cooperation in the prisoner's dilemma. Specifically, cooperation was more likely for individuals who scored higher on the withdrawal aspect of neuroticism and the enthusiasm aspect of extraversion. Withdrawal is associated with fear and insecurity, suggesting that individuals who score highly on this trait may cooperate because they are afraid of the consequences of defecting. Indeed, behavioral economists have posited that fear of punishment can play a substantial role in the promotion of cooperative behavior (Fehr & Gächter, 2000; Hirshleifer & Rasmusen, 1989). Enthusiasm, meanwhile, is the aspect of extraversion that is associated with positive affect and sociability. Individuals who score highly on this trait tend to experience more positive emotion and are more sensitive to reward (Depue & Collins, 1999). Enthusiastic people may be experiencing a greater subjective reward from cooperative behavior, increasing the chance of subsequent cooper-

**Table 1**  
Descriptive statistics.

	M	SD
Extraversion	3.39	0.47
Agreeableness	3.78	0.46
Conscientiousness	3.18	0.41
Neuroticism	3.00	0.56
Openness/intellect	3.50	0.47
WPT	21.40	5.62
Years English	13.87	6.40

ation (cf. Andreoni, 1990; Rilling et al., 2002). Positive affect may also be playing an important role in the prisoner's dilemma, such that individuals in positive moods are more likely to experience cooperation as rewarding, and to expect their partners to cooperate. Interestingly, the two personality findings that emerged from the current analysis suggest that there may be multiple pathways to cooperation in the prisoner's dilemma: one based on fear of punishment, and the other based on the rewarding aspects of cooperation (cf. Andreoni & Harbaugh, 2003). Future research can examine this possibility by examining whether situational threats or rewards are more effective inducers of cooperation for individuals with different personality traits. In particular, it is expected that individuals higher in neuroticism would cooperate more in response to threats, whereas extraverts would cooperate more in response to rewards (Carver, Sutton, & Scheier, 2000).

Overall, the study suggests that personality can be a significant predictor of behavior in the prisoner's dilemma. More generally, it suggests that personality psychology can be an effective framework from which to understand individual differences in decision-making and economic behavior. The study could be improved upon by employing a more gender-balanced sample and examining causal, rather than correlational effects. Given the emotional correlates of both withdrawal and enthusiasm, future research should examine whether cooperation in the prisoner's dilemma would also be observed following situational inductions of positive or negative emotion.

## References

- Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *Economic Journal*, *100*(401), 464–477.
- Andreoni, J., & Harbaugh, W. (2003). The carrot or the stick: Rewards, punishments, and cooperation. *American Economic Review*, *93*(3), 893–902.
- Axelrod, R., & Hamilton, W. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390.
- Bendor, J., Kramer, R., & Stout, S. (1991). When in doubt: Cooperation in a noisy prisoners' dilemma. *Journal of Conflict Resolution*, *35*(4), 691–719.
- Boone, C., De Brabander, B., & van Witteloostuijn, A. (1999). The impact of personality on behavior in five prisoner's dilemma games. *Journal of Economic Psychology*, *20*(3), 343–377.
- Carver, C. S., Sutton, S. K., & Scheier, M. F. (2000). Action, emotion, and personality: Emerging conceptual integration. *Personality and Social Psychology Bulletin*, *26*(6), 741.
- Cooper, R., DeJong, D., Forsythe, R., & Ross, T. (1996). Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, *12*(2), 187–218.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory and NEO five-factor inventory professional manual*. Odessa, FL: Psychological Assessment Resources.
- Davis, D. D., & Holt, C. A. (1993). *Experimental economics*. Princeton: Princeton University Press.
- Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences*, *22*(3), 491–517.
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: Ten aspects of the big five. *Journal of Personality and Social Psychology*, *93*(5), 880–896.
- Dodrill, C. B. (1981). An economical method for the evaluation of general intelligence in adults. *Journal of Consulting and Clinical Psychology*, *49*, 668–673.
- Dodrill, C. B., & Warner, M. H. (1988). Further studies of the Wonderlic personnel test as a brief measure of intelligence. *Journal of Consulting and Clinical Psychology*, *56*, 145–147.
- Ellison, G. (1994). Cooperation in the prisoner's dilemma with anonymous random matching. *Review of Economic Studies*, *61*(3), 567–588.
- Fehr, E., & Gächter, S. (2000). Cooperation and punishment. *American Economic Review*, *90*(4), 980–994.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*(1), 26–34.
- Hirshleifer, D., & Rasmusen, E. (1989). Cooperation in a repeated prisoners' dilemma with ostracism. *Journal of Economic Behavior and Organization*, *12*(1), 87–106.
- John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 102–138). New York: Guilford Press.
- Killingback, T., Doebeli, M., & Knowlton, N. (1999). Variable investment, the continuous prisoner's dilemma, and the origin of cooperation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1430), 1723–1728.
- Kreps, D., Milgrom, P., Roberts, J., & Wilson, R. (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, *27*(2), 245–252.
- Kuhlman, D. M., & Marshello, A. F. (1975). Individual differences in game motivation as moderators of preprogrammed strategy effects in prisoner's dilemma. *Journal of Personality and Social Psychology*, *32*(5), 922–931.
- McKelvie, S. J. (1989). The Wonderlic personnel test: Reliability and validity in an academic setting. *Psychological Reports*, *65*, 161–162.
- Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, *364*(6432), 56–58.
- Rapoport, A., & Chammah, A. M. (1965). *Prisoner's dilemma*. Ann Arbor, MI: University of Michigan Press.
- Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron*, *35*(2), 395–405.
- Wonderlic, E. F. (1983). *Wonderlic personnel test manual*. Northfield, IL: E.F. Wonderlic & Associates.