ELSEVIER

Journal of Economic Psychology 23 (2002) 49–66

JOURNAL OF Economic Psychology

www.elsevier.com/locate/joep

Using the Machiavellianism instrument to predict trustworthiness in a bargaining game

Anna Gunnthorsdottir [a,*], Kevin McCabe [b], Vernon Smith [b]

[a] *Department of Management and Policy, University of Arizona, McClelland Hall 405, Tucson, AZ 85721-0108, USA*

[b] *Economic Science Laboratory, University of Arizona, Tucson, AZ, USA*

Received 22 December 1999; received in revised form 1 August 2001; accepted 3 September 2001

Abstract

Game-theoretic experiments have revealed substantial individual differences where the game allows for off-equilibrium behavior such as trust and reciprocity. We explore the personality psychology and decision making literatures and conclude that these individual differences are likely to be mediated by differential emotional arousal. We argue that Christie and Geis's Machiavellianism scale (Mach-IV) is an instrument that allows the identification of types who vary in cooperativeness. We use that test to predict the behavior of participants in a two-person one-shot constituent game in which subjects face a choice between trust and distrust, and between reciprocation (trustworthiness) and defection. We find that the Mach-IV scale does not predict trusting behavior. It does, however, predict reciprocity. Over one half of those who score low to average on the scale reciprocate trust. High scorers overwhelmingly defect when it is to their advantage to do so.

PsycINFO classification: 3020; 3120

JEL classification: C72

Keywords: Game; Machiavellianism; Reciprocity; Trust

* Corresponding author. Tel.: +1-520-621-1053; fax: +1-520-577-8389.
E-mail address: agunntho@hotmail.com (A. Gunnthorsdottir).

0167-4870/02/$ - see front matter
PII: S0167-4870(01)00067-8

1. Introduction

1.1. Trust, reciprocity and individual differences

Trust is a voluntary transfer of a good or favor to someone else, with future reciprocation expected but not guaranteed. Reciprocity is voluntarily repaying a trusting move at a later point in time, although defaulting on such repayment is in the short-term self-interest of the reciprocator. If trust is reciprocated both parties gain from the exchange.

Humans take trust and reciprocity for granted and often fail to notice the pervasiveness of these behaviors. A propensity to exchange favors and goods is observed cross-culturally and such acts are ubiquitous in interactions between friends, colleagues and acquaintances. Examples include children lending each other crayons, hunter–gatherers sharing meat, and verbal commitments between political negotiators (Barkow, Cosmides, & Tooby, 1992; Brown, 1991, pp. 137–338; Kaplan & Hill, 1985; Lykken, 1995).

Such cooperation is not necessarily in conflict with game-theoretic predictions of what self-interested agents should do. An other-regarding act, even if not reciprocated, often enhances one's reputation. Moreover, social interactions are usually of indefinite length. There is most often a chance, however small, to meet again. This makes the interaction a game of infinite length. In such games many equilibria are possible, cooperation included. However, the subgame perfect solution fails to account for the ubiquity of cooperation among humans; it only allows for it as one of many possible equilibria.

More importantly, trust and reciprocity persist in game-theoretic experiments in which interactions are anonymous and finite and when therefore no such behaviors are expected as long as we assume that only material payoffs count. For example, in dictator games subjects show off-equilibrium other-regarding behavior when they voluntarily and to varying degrees share a fixed sum of money with their counterpart (Forsythe, Horowitz, Savin, & Sefton, 1994). A similar propensity toward equitable outcomes rather than using one's full advantage has been observed in multi-stage bargaining games and again individual differences are very apparent (Thaler, 1992). Non-material motives in experimental games have been proposed and examined (see for example Bolton, 1991; Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999; Gunnthorsdottir, Houser, McCabe, & Ameden, 2001; Palfrey & Prisbrey, 1997; Rabin, 1993) and this research, too, has revealed substantial individual differences in preferences. Not only do preferences vary when the experimental environment is kept invariant, but while anonymity generally weakens other-regarding behavior, subjects also react differentially to variations in experimental anonymity itself (Hoffman, McCabe, & Smith, 1996a). It thus appears that there is variation both in individuals' reciprocity norms and in the extent to which such norms are ingrained in them.

1.2. Extensive-form bargaining games

McCabe, Rassenti, and Smith (1996) let subjects interact anonymously in extensive-form bargaining games in which anonymous pairs of players alternated their

moves along a game tree, with options of trust, non-trust, reciprocity and defection. Some 46% of players trusted in a one-shot game. Half of their counterparts reciprocated this trust to reach a mutually advantageous outcome. In fact, given the frequency of reciprocation the expected return from trust was as high as that of non-trust.

Berg, Dickhaut, and McCabe (1995) examined a one-shot two-person investment game with continuous strategy space in which a first mover could entrust his or her counterpart with any fraction of his or her $10 show-up fee. Any amount passed on was known to triple before the second mover received it. The second mover then had to decide how much of the money, if any, to return to the first mover. The second mover was in a dictator position but, in contrast to a traditional dictator game, the money he controlled did not come from the experimenter but from a trusting counterpart whose investment had generated substantial gains from exchange. Thirty of 32 first movers entrusted their anonymous counterpart with at least some of their $10 show-up fee. The exact amounts were very variable. With regard to reciprocity, too, there were substantial individual differences. While some second movers kept all of the money, a substantial proportion reciprocated sufficiently for both players to make a profit.

1.3. The $10 Trust Game

The $10 Trust Game shown in Fig. 1 is a simple bargaining game in extensive form and a truncated version of the extensive-form game devised by Berg et al. (1995).[1] Numbers in decision nodes reflect temporal order as well as which player moves at that node. Payoffs are shown in the boxes at each endpoint, with a Player 1's payoff above Player 2's. Players alternate in moving along the decision tree by selecting between the branches departing from their decision node. Player 1 moves first and faces two options. He or she can move right, avoiding interaction and ending the game. In this case each participant receives $10. A riskier but potentially more rewarding move is for Player 1 to move down and give Player 2 a chance to move. In this case whatever Player 2 does, joint payoffs double (from $20 to $40) representing the gains from cooperation. Player 2 now faces two options: he or she can defect by moving down, take the entire joint payoff of $40 and leave Player 1 with nothing. Alternatively, he or she can move right. Such a right move by Player 2 gives both players more than if Player 1 had moved right and ended the game ($15 vs. $10 for Player 1, $25 vs. $10 for Player 2). Player 1 signals to his or her anonymous counterpart through his or her very move. Moving right conveys distrust in Player 2. Moving down signals trust and a desire to arrive at the mutually advantageous outcome of [15, 25].

[1] A move down by Player 1 in the $10 Trust Game corresponds to an investment of his or her $10 show-up fee, which, on the way to Player 2, triples to $30. If Player 2 moves right, this corresponds to splitting the $30 evenly, yielding and end result of $15 for Player 1 and $25 for Player 2 ($15 plus Player 2's show-up fee of $10). If Player 2 moves down, this corresponds to him keeping the entire $30.

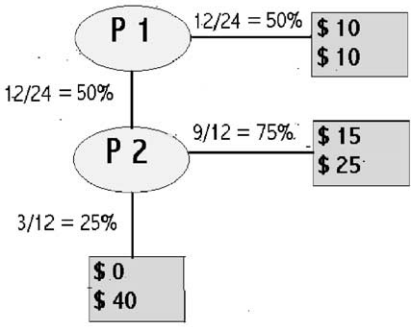

Fig. 1. Frequencies of trust and reciprocity in the $10 Trust Game – unscreened subjects.

The noncooperative equilibrium in this two-stage game, arrived at by backward induction, is so easy to determine that it is hard to imagine that any subject might not fully grasp the risks of deviating from it. Camerer, Johnson, Rymon, and Sen (1993) find in a shrinking pie game that all subjects look ahead one stage, and based on their results one would expect the same here. Starting with Player 2 at node 2, if he or she is rational and self-interested, he or she moves down to get $40. Player 1 can anticipate this and therefore will move right. The equilibrium outcome $[10, 10]$ is less efficient than the $[15, 25]$ players could attain through off-equilibrium trust and reciprocity. When the $10 Trust Game was played by 24 unscreened undergraduate students who interacted once and anonymously, subjects exhibited the behavioral heterogeneity so frequently observed in experimental bargaining games. Twelve of 24 first movers trusted their counterpart and moved down. Their trust was rewarded by 75% of their counterparts who reciprocated by moving right (see Fig. 1). First movers who moved down made on average somewhat more than those who moved right ($11.25 vs. $10) (McCabe & Smith, 2000).

1.4. Why individual differences in trust and reciprocity?

While standard economic theory cannot account for the observed variability in trust and reciprocity, social psychologists have found personality scales helpful in predicting behavioral variance. From a different perspective, evolutionary psychologists have focused on how the existence of cooperators (individuals who trust and reciprocate) supports the emergence of non-reciprocators or "cheaters", who consistently tend to take advantage of extended trust. Although we make no attempt to review this literature here, we note in particular the work of Trivers (1971), Barkow et al. (1992; see also Tooby & Cosmides, 1995), Gigerenzer (1996), Cummins (1999), Brown and Moore (2000), and Wedekind and Milinski (2000).

A further argument supporting the existence of types in interpersonal behavior is the fact that boundedly rational humans are unable to respond optimally to each sin-

gle situation. Rather, each individual develops his or her own set of behavioral strategies that are then applied across situations. Such decision heuristics allow speedy response but do not permit precise adaptation to specific situations (Grusec, Lockhart, & Walters, 1990). If heuristics such as when and to what extent to trust and reciprocate vary among individuals, their existence should manifest itself as types, which vary in the amount of trust they show and in the extent of "cheating" they engage in.[2]

Heuristics are often emotionally based. Emotions tend to process the information inherent in a decision situation before it has been consciously grasped, causing a semi-automatic response (Damasio, 1994). Lykken (1995) argues that interpersonal decision-making involving values, such as the repayment of trust, is often executed via such emotionally based heuristics, which people acquire through classical conditioning in childhood. Due to genetically based individual differences in emotional arousability, individuals vary in the ease and speed with which they respond to conditioning. Children with cool affect are less likely to internalize social norms including reciprocity, and will end up with relatively weak internal standards, with what is commonly called "a weak conscience". For example, they fail to reciprocate when it is more advantageous not to. Hirshleifer (1987) and Frank (1988) have made similar arguments about the role of affect as a commitment device for behavior that would be superficially deemed irrational. Cool affect is therefore a characteristic we would expect in a cheater, who is not bound by social norms across situations but rather uses whatever means he or she deems appropriate in order to achieve an end.

Mealy (1995) argues that sociopaths are typical "cheaters" i.e. non-reciprocators. Sociopathy is a personality disorder and thus a distinct clinical category. While possessing superficial charm, sociopaths are unreliable, untruthful, exploitative, lack remorse and shame, and engage in antisocial behavior. Physiological measures support the common perception that sociopaths lack emotion (Cleckley, 1955; Hare & Craigen, 1974; Hare, 1993; Lykken, 1995). However, Mealy notes that disregard for social norms is not limited to sociopaths. Normals certainly vary in their commitment to conventional morality and in the robustness of their reciprocity strategies in the face of temptation. Mealy (1995, p. 534) suggests that the personality dimension of Machiavellianism could serve as a measure of subclinical tendencies to cheat.

1.5. Machiavellianism

When identifying cheaters, Machiavellianism has the advantage over sociopathy that it is not a distinct clinical category but rather a personality trait measured along a continuum. Christie and Geis (see Christie, 1970a) first drew attention to this trait

[2] An interesting suggestion was made by an anonymous reviewer, namely that the existence of types could affect learning models such as the Erev and Roth (1998), Roth and Erev (1995) reinforcement model. Clearly, if types could be distinguished at the outset and individual differences in preferences incorporated into the model, a learning model's predictive power should increase.

and developed an instrument to measure it. Their Mach test was based directly on Machiavelli's treatises "The Prince" and "Discourses on the First Decade of Titus Livius", and items were chosen to reflect Machiavelli's general views and recommendations. The scale was designed to measure a test-taker's tendency toward Machiavellian behavior, in particular cynicism, a penchant for manipulativeness, and the belief that the end sanctifies the means.

The Mach scale has since been used in more than 500 psychological studies, both experimental and demographic. In the 1970s and 1980s many experiments focused on validating the predictive power of the test by assessing how the behavior of high and low Machs differs. With the test's predictive validity eventually a given (Fehr, Samson, & Paulhus, 1992), there has been some decline in studies assessing behavioral differences between high and low scorers. [3] However, recently there has been renewed interest in the test due to the emergence of the "cheater detection" literature, as the Mach test appears to be an instrument to identify such non-reciprocators. In a comprehensive overview of the intersection of personality literature and evolutionary literature Wilson, Near, and Miller (1996) advocate the use of Machiavellianism scales to test evolutionary cooperator–defector models, and suggest an array of testable hypotheses derived from evolutionary theory. [4]

The general conclusion from the numerous experimental comparisons of high Mach vs. low Mach behavior is that even though Machiavellianism is not correlated with IQ (Wilson et al., 1996), Highs outperform Lows in most of the short-term interactions of typical experiments, especially so if an environment allows for face to face contact, latitude for improvization, and arousal of irrelevant affect. In spite of the impressive exploits of high Machs in short-run encounters, however, Lows seem to catch up somehow in the long term and Mach scores, for example, are not correlated with social status or income (Fehr et al., 1992). What follows is an introduction of experimental findings relevant to our hypotheses about how Machiavellianism might affect behavior in the $10 Trust Game (see Fig. 1).

1.5.1. Means sanctifying the ends

High Machs have a general cool attitude, manifested as a detached, opportunistic stance toward values social norms (Mudrack & Mason, 1995). Rational and often materialistic, Highs can calmly identify the optimal strategy in each situation (Christie & Geis, 1970) and behave in a self-interested fashion if it is to their advantage (Effler, 1983). They are more upset by inefficiency than by injustice (Christie & Geis, 1970, p. 353) and thus appear to be true *homines economici* who continuously test

[3] According the PsychInfo database there were only about 125 published studies that used the test in the last decade compared to about twice that in each of the two preceding decades. More recent psychological studies of Machiavellianism have tended to use the Mach test in order to validate more recently developed tests.

[4] They convincingly argue that evolutionary models would benefit from empirical verification, which requires identifying cooperator types. Studies in Machiavellianism on the other hand, have hitherto lacked a theoretic framework which evolutionary theory can provide.

how far they can go. Highs are "called 'opportunistic' by those who deplore [their behavior], [but] 'realistic' by more admiring observers" (Christie & Geis, 1970, p. 303). Note, however, that unlike sociopaths who are often marginal, high Machs are usually well integrated into society.

Highs are more apt than Lows to behave unethically (Jones & Kavanagh, 1996). While high Machs are not distracted by affect, Lows cannot easily separate themselves from moral precepts in situations that offer material rewards for breaking norms (Geis & Christie, 1970). They therefore tend to adhere to implicit and explicit rules such as reciprocity (Christie & Geis, 1970, p. 351). In the same vein, social pressure affects Lows more than Highs. For example, Highs cheat if given rational arguments in favor of it, and when the probability of detection is low. Low Machs cheat, too, but mostly if coaxed into it by someone else (Cooper & Peterson, 1981; Exline, Thibaut, Hickey, & Gumpert, 1970).

1.5.2. Manipulativeness

Highs are both more manipulative (Cherulnik, Way, Ames, & Hutto, 1981; Wilson, Near, & Miller, 1998) and more exploitative than Lows (Vecchio & Sussmann, 1991). They tend to lie more (Znakov, 1995) and more convincingly (Exline et al., 1970; Geis & Moon, 1981). Lows, in contrast, get emotionally involved with people they encounter. For example, Highs are more ready than Lows to leave an implicit alliance when it is to their advantage (Geis & Christie, 1970), and while low Machs are less likely to steal from someone if that person trusts them, Highs take advantage of extended trust (Harrell & Hartnagel, 1976). Nonetheless, Highs are generally perceived as more likeable than Lows (Wilson et al., 1998).

1.5.3. Cynicism

High Machs are generally more likely to suspect others of dishonesty (Harrell, 1980). Already the original creators of the Mach scales noticed that their high Mach subjects were less credulous than Lows with regard to experimental manipulations (Geis & Christie, 1970). These findings about high Machs being both exploitative and cynical echo results obtained by social dilemma researchers that non-cooperators in Prisoner's Dilemma games also expect less cooperation from others (e.g. Orbell & Dawes, 1993).

However, high Machs' behavior might appear more trusting than their cynicism warrants, because Highs may be more risk-seeking than Lows. Trusting involves taking a risk in an interpersonal encounter. For example, Boone, De Brabander, and van Witteloostuiijn (1999) found that cooperation in Prisoner's Dilemma games is correlated with Zuckerman's scale of sensation-seeking. Like high sensation-seekers sociopaths, too, are known to seek risk due to hypoarousal (Daderman, 1999; Eysenck & Gudjonsson, 1989). High Machs, as subclinical sociopaths (McHoskey, Worzel, & Szyarto, 1998), might therefore be willing to take chances in interpersonal encounters – a behavior that may be mistaken for a trusting interpersonal orientation. The few studies that have dealt with the relationship between Machiavellianism and risk-attitude (Allsopp, Eysenck, & Eysenck, 1991; Tamborni, Stiff, & Zillmann, 1988) indicate that high Machs may indeed be more risk-seeking than Lows.

1.6. A cautionary note on predicting strategic behavior using personality measures

It would be empirically useful and fully compatible with game-theoretic concepts, if we could find a simple scoring task that identified player types; in particular, types that would differ in their propensities to be trusting (cooperative) or be trustworthy (reciprocators). To be most useful as a prescreen in game-theoretic experiments, a personality measure must predict a small number of behavioral choices in specific situations and for modest sample sizes. Yet in single situations personality measures tend to show a low R^2 with behavior they are expected to predict. Personality measures identify general behavior patterns, and their predictive power increases with aggregate observations over time and situations (Epstein, 1979). The power of personality tests in predicting behavior in experiments may therefore be weak. (For a critique of cultural theory based on personality measures, see Sjöberg, 1997.)

It should be noted, however, that a low R^2 does not necessarily imply that a predictor is of little importance. Thus, Rosenthal (1990) provides several examples in psychology and medicine where treatment effects are of great practical importance, but the R^2 were as low as 11% (effect of aspirin on heart attacks). For example, treating AIDS patients with AZT reduces deaths from 61.5% to 38.5%; dramatic enough to cause termination of the trials, although the R^2 between survival and treatment with AZT was only 5% (see also Slovic & Peters, 1998). Similarly, increasing the likelihood of detecting (and consequently avoiding) a non-reciprocator, however small the increase, may substantially increase the expected gains from exchange, illustrating the difference between statistical and economic significance.

1.7. Machiavellianism and experimental games

High Machs with their cool, rational attitude should be true gamesmen, and better than Lows at going after their short-term self-interest in anonymous and finite interactions. While Mach scales have been used extensively in psychological research, they have not been widely used in the context of experimental games. The few experiments of this type, mainly from the 1960s, have focused on the Prisoner's Dilemma (e.g. Ueijo & Wrightsman, 1967; Wrightsman, 1966), related normal-form pairwise interactions (Christie, Gergen, & Marlowe, 1970), and the game of Chicken (Marin, 1973), and have yielded mixed results (see Terhune, 1970, for an overview of early game-theoretic experiments involving the Mach test. There have been, to our knowledge, no other published experiments using the Mach test in the context of games until the present study with the exception of Meyer, 1992 discussed below).

Wilson et al. (1996) suggest that this apparent lack of results is due to experimenters failing to compare high and low Mach behavior over repeated trials. According to them, high Machs should be responsive to the strategic distinctions between one-shot and repeated games. Highs should defect in one-shot games but reciprocate in repeat interaction at least initially, in order to induce their counterpart to engage in mutually advantageous cooperation. Low Machs, on the other hand, should remain

cooperative in both situations. This hypothesis has yet to be tested with regard to positive reciprocity. However, Meyer (1992) confirms it with regard to negative reciprocity (retaliation). He found low Machs more willing to refuse low offers as second movers in an ultimatum game, both in single and in repeated play, supposedly due to inherent fairness norms. Highs, on the other hand, were willing to accept low offers in single play but rejected them in repeated interaction. Presumably, by refusing low offers in repeated interaction, Highs were trying to force their counterpart to better his or her offers.

Further, and related to Wilson et al.'s (1996) argument of why high and low Machs have not hitherto been found to consistently differ in experimental games, we should like to point out that in the games used so far, such as the two-person Prisoner's Dilemma, players move simultaneously. This gives rise to behavioral ambiguity. In a Prisoner's Dilemma game, a noncooperative move may either signal lack of trust, an intention to take advantage of the other person, or both. Neither a player nor an outside observer can clearly infer behavioral intent from such a move. Simultaneous-move games thus fail to capture the sequential interactions of reciprocal exchange in which a trusting move is interpreted as a signal of cooperative intent, and is either reciprocated or exploited by the other player (McCabe, Smith, & LePore, 2000; see also Kiyonari, Tanida, & Yamagishi, 2000). Differences in the behavior of high and low Machs in one-shot games are more likely to emerge in extensive-form bargaining games such as the $10 Trust Game (see Fig. 1), where players alternate in their moves and sequentially choose between trust and non-trust, and between reciprocity and defection.

1.8. Hypotheses

How should high Machs differ from low Machs when playing the $10 Trust Game (see Fig. 1)? The first move involves a choice between trust and non-trust, and two contradictory predictions can be derived. As Highs are more cynical than Lows (Geis & Levy, 1970) and because generally in games, non-cooperators expect others to be like them (see, e.g. Orbell & Dawes, 1993), Highs should move right more often than Lows. On the other hand, if Highs are indeed more risk-seeking than Lows, it should motivate them to make the risky move down. The authors therefore concluded a priori that the Mach score is ambiguous in predicting the behavior of Player 1. Any test of behavioral differences between high and low Machs in the Player 1 role must therefore be two-tailed.

The second move involves a choice between reciprocation (by choosing $[15, 25]$) and taking advantage of Player 1's trust (by choosing $[0, 40]$). Here predictions are clear for high Machs: a typical cheater abuses trust and takes all he or she can. As high Machs easily separate themselves from reciprocity norms and can size up an interaction in a detached, rational manner, we expect them to play the dominant strategy and choose the outcome that gives them $40 even if it leaves their counterpart empty-handed. Low Machs, on the other hand, cannot separate themselves as easily from reciprocity norms and get emotionally involved in the interaction with their counterpart. They should therefore reciprocate and move right.

2. Method

2.1. Measures

To assess Machiavellianism, we used the Mach-IV scale developed by Christie and Geis. This paper-and-pencil test consists of 20 statements in three substantive areas: Machiavellian views, tactics and morality (see Appendix A). Raw scores rank from 20 to 140. A constant of 20 is added to all raw scores so that the scale ranges from 40 to 160, with a theoretical midpoint of 100. The higher the score, the stronger the test-taker's Machiavellian orientation.

2.2. Reliability

To our knowledge there have been no studies assessing the stability of Machiavellianism over an extended period. Meyer (1992) reports a test–retest reliability of 0.73 over two-weeks ($N = 126$). As our entire research project on Machiavellianism extended over nearly nine months, we assessed the test–retest reliability of the Mach-IV over this interval. During that time, test-takers nearly completed their freshman year and participated in at least one game-theoretic experiment. The correlation between their initial test score and their score nine months later was a satisfactory 0.64 [5] indicating Machiavellianism's temporal stability.

2.3. Subjects

1593 freshmen taking an introductory psychology class took the Mach-IV test (see Appendix A) either in September 1997 or in January 1998, as part of a large battery of tests administered by the Psychology Department. Women were slightly over-represented among test-takers (61.17%). To our knowledge, such a large group has not been tested for Machiavellianism since the 1960s when Christie (1970b) used samples of similar size.

Table 1 reports the descriptive statistics for our sample and compares it to the 1964 college sample of Christie (1970b). Consistent with previous research (Christie, 1970b; Gupta, 1987) women turned out to be less Machiavellian than men: Table 1 shows that our statistics generally resemble Christie's, but the average Mach score in our sample was higher. This is consistent with previous conjectures (Christie, 1970b;

[5] Another 15 students took the Mach-IV twice within a five-month interval and their correlation was 0.87. We deem $r = 0.64$ satisfactory because of the long test–retest interval. For very short intervals a test–retest reliability of $r = 0.8$–0.9 is desirable, and progressively decreases as the test interval lengthens (Anastasi, 1990, p. 117). There are two reasons why test–retest correlations are never unity: first, unreliability of the test per se, and second, and increasingly important as the test interval lengthens, actual change in the trait over time.

Table 1
Mach-IV scores in our sample; compared to Christie's 1964 sample of college students (S.D., where available, are in parentheses)

	University of Arizona	1964 college sample
N	1593	1744
Median age	18	19
% Female	61.2	51.7
Average score females	90.77 (12.3)	87.69
Average score males	96.69 (13.5)	93.94
Average score whole sample	93.00 (13.1)	90.12[a] (14.33)

[a] Weighted to match male–female proportions in Arizona sample.

Wrightsman & Wuscher, 1974, p. 379) that Mach scores of Americans are increasing with subsequent cohorts.[6]

From our tested freshmen population, we recruited 266 individuals who scored either low ($\leqslant 80$, up to the 16th percentile), average-to-low (81–90, percentiles 16–41),[7] average (91–95, percentiles 42–58), or high ($\geqslant 107$, 88th percentile and above) on the Mach scale. As the Mach test had been administered as part of a large test battery, by another department, and in a different location, it was practically impossible for students to guess that a specific part of the test battery was connected to a later invitation to a paid economics experiment. In order to enlarge the subject pool and to ease recruiting Mach scores were collected during two semesters, in late August 1997 ($N = 1132$) and in January 1998 ($N = 461$). Students who took the Psychology Department's test battery in January 1998 were added to the subject pool at that time. The experiments themselves were conducted between late November 1997 and February 1998.

2.4. Experimental design and procedure

Subjects were recruited for a one-hour experiment with a $5 show-up fee and the possibility of additional earnings depending on their decisions. The experimental design was balanced, with three pairings and six subject pairs per session: (1) two Low–High pairs (with a low Mach as the first mover, a high Mach as the second mover), (2) two Low–Low pairs, and (3) two pairs in which both players had average Mach scores. After 19 such sessions the subject pool was not yet exhausted. We therefore ran five additional sessions in which high Machs were in the first mover position[8] in

[6] An alternative explanation for a change in Mach scores over cohorts, as pointed out by an anonymous reviewer, may be some Mach-test items are not gender-neutral (Items 14, 18 and 20, see Appendix A). However, it is worth noting that, in spite of changing views about gender, the gender gap in Machiavellianism has remained the same over nearly four decades.

[7] Originally, we had planned to use only low, average and high Machs. Average-Lows were used when it was decided to run further experiments and more subjects were needed.

[8] In this latter series of experiments, nine high Machs who had previously participated as second movers were allowed in the first mover position. During their first participation they had not got a chance to move at all, because Player 1 in their dyad had immediately moved right and ended the game.

order to examine possible differences in trust between Lows and Highs. In these five sessions the second mover position was filled with subjects who ranged from low to average in Machiavellianism. Since subjects were costly and difficult to recruit, we overbooked all categories of Mach scores and we occasionally admitted an additional pair. We also ran sessions with as few as four pairs, and if we could not form exactly the pair combination planned the remaining subjects were arranged in one of the other treatment pairs.[9]

We followed the standard procedures at the University of Arizona's Economic Science Laboratory. Each subject revived $5 in cash upon arriving in time for the experiment. Immediately upon arrival, subjects were randomly assigned to a computer terminal reserved for their type and seated at their terminal. This minimized any socialization among waiting subjects. The terminals were separated by blinders and subjects never sat at adjacent terminals. Participants privately went through computerized instructions with practice examples of game trees, including alternating moves with a simulated counterpart. At the start of the actual experiment, an experimenter reiterated that subjects would interact once with an anonymous counterpart in the room, with payoffs for both members of a dyad displayed on both their screens, and a subject's own potential dollar payoffs highlighted. Subjects then played the $10 Trust Game shown in Fig. 1 for the dollar amounts shown. At the end of the experiment participants filled out a short questionnaire and were then privately paid any experimental earnings.

3. Results

As males, on average, score higher on the Mach test than females, men were slightly over represented among our high Machs compared to the proportion of men in the whole sample. Similarly, women were over represented among low Machs. However, no trace of a gender effect on behavior remained once Mach scores had been taken into account.

Since there were no significant differences in behavior between low, average-to-low and average Machs, their data were combined to form the average-low group (AL). Results are summarized in Fig. 2. It shows frequencies and percentages of moves separately for high (H) and average-low (AL) Machs in each mover position. There is no statistically significant difference among types in the frequencies with which they trust and move down as Player 1. In the Player 2 position, however, we found that while well over half of low to average Machs reciprocated trust and

[9] In eight sessions an odd number of subjects showed up and one of the experimenters assumed the first mover position against the odd subject, forming an additional experimenter–subject pair. The experimenter always moved down. On one further occasion, the experimenter was in the second mover position. We take care to avoid deception in experiments because of its potential for contaminating the subject pool if practiced as a matter of policy. In this study, 5% of our subjects interacted with an experimenter rather than a fellow-subject and were not informed that their counterpart was not another subject. Data from three subjects were omitted as they turned out not to be in our pre-specified subject pool.

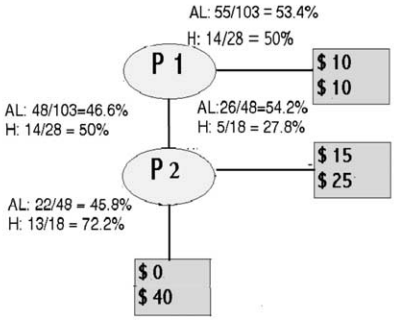

Fig. 2. Frequencies and percentages of trust and reciprocity, by type, for high (H), and average to low (AL) Mach scores.

moved right, over two thirds of high Machs in the Player 2 position played the dominant strategy and defected. A one-tailed binomial test shows that the differences in frequency in the second mover position are significant at less than the 0.05 level ($U = 1.64,\ p = 0.03$).

Comparing Figs. 1 and 2, we also tested whether the proportion of unscreened, regular subjects who reciprocated (9/12) differs significantly from the proportion of average-scoring Machs who reciprocated (12/20). This difference is not significant and we conclude that our unscreened subjects exhibit reciprocity behavior that is comparable to the average Machs in our sample.

4. Discussion

Lykken (1995) and Frank (1988), among others, have suggested that individual differences in reciprocity are due to differential emotional arousability. Mach scales help identify individuals with a cool, rational attitude. We have argued that measurements of Machiavellianism should therefore allow us establish individuals' propensity to reciprocate or to defect in an anonymous, one-shot experimental game where defection is the equilibrium strategy.

In the past, the use of personality measures to account for behavioral variance in game-theoretic experiments has been very limited, and for all personality measures studied the results have been mixed. This might be due to the fact that in the past personality measures have been matched to games rather indiscriminately.

First, situations may be either "strong", making a strong prescription for certain behavior independent of personality type, or "weak", allowing personality to affect behavioral choices (Mischel, 1977). A counterpart's behavior is probably a strong general influence on behavioral choices (see, for example, Clark & Sefton, 2001; Hayashi, Ostrom, Walker, & Yamagishi, 1999), even more so in repeated games. For example, in a game repeated several times, repeated defection of one side eventually calls for defection by the other side as well. In order to assess the effects of personality in games, a gaming situation must allow leeway for personality to assert

itself. For this reason, unless specific hypotheses about strategic behavior in repeated games are tested (such as the ones suggested by Wilson et al., 1996), one-shot games may generally be better suited for assessing the effect of personality in games.

Second, a personality measure cannot account for behavior in a game if the trait it measures is not relevant to the situation at hand. Matching a game to traits that really matter in that situation is a prerequisite for getting results. In the present study, especially for Player 2, different behavioral choices were to be expected for high vs. low Machs. Machiavellianism is clearly relevant to Player 2's decision task, which involves social norms and a choice between responding to a cooperative signal or ignoring that signal and realizing one's self-interest to the detriment of one's counterpart.

Even in situations where the traits are well matched to decision tasks at hand, however, a measure of a personality trait such as Machiavellianism merely predicts differences in behavior among aggregates of types. Even if a trait measure were perfectly reliable and valid for the situation at hand, some behavioral variation within types would remain due to individuals' varying states and differences in how they frame and interpret a situation (see for example Hoffman et al., 1996a,b). Due to these complexities personality tests typically only allow to generate hypotheses about the existence and direction of effects, rather than effect magnitude. These considerations hold for this study as well.

Highs' and Lows' behavior apparently does not differ in the first mover position where the task is reading the mind of the second mover in order to assess whether he or she is trustworthy. Statistically, all types appear equally divided between a trusting and a non-trusting move. However, it remains to be established whether the same motivational and cognitive processes underlie these identical decision frequencies. The null results for Player 1 do not unequivocally support our original ambivalence about forming a directional hypothesis about how Mach scores should predict first mover decisions. The results are nonetheless in agreement with our initial conjectures about conflicting motives in that position. We had speculated that high Machs may be less risk-averse but also less trusting, and Lows more trusting but also more risk-averse. Further experimentation with separate measures of trust and risk attitude could clarify whether Mach scores indeed fail to predict Player 1 behavior because Player 1 must strike a balance between such conflicting motives.

For second movers, our data unequivocally support our hypothesis that high Machs reciprocate less than others in the population do. The Machiavellianism scale is therefore useful as an instrument to detect individual tendencies toward reciprocity and defection. In other words, the test can assist with cheater detection. Highs tend to not reciprocate in an anonymous situation where they can get away with such behavior. As recommended by Machiavelli in his 16th century treatises, they tend to use morally questionable means (not repaying extended trust) to achieve their ends (maximum earnings). Long before the development of psychological testing, individuals and society have developed mechanisms to counteract such behavior by increasing the likelihood of identifying cheaters who, in the long run, reduce social capital and efficiency (Coleman, 1990). Such mechanisms range from cognitive structures specific to this important task such as cheater detection modules (Barkow et al., 1992; Gigerenzer & Hug, 1992), and mental goodwill accounting systems (Smith &

McCabe, 2001) to gossip (Dunbar, 1996) and public or semi-public records. In an increasingly mobile and anonymous society systematic psychological testing such as use of the Mach scale can serve as an instrument to screen a larger population for cheater tendencies.

Appendix A. Mach-IV scale

1. Never tell anyone the real reason you did something unless it is useful to do so.

strongly agree	somewhat agree	slightly agree	no opinion	slightly disagree	somewhat disagree	strongly disagree

2. The best way to handle people is to tell them what they want to hear.
3. One should take action only when sure it is morally right. [10]
4. Most people are basically good and kind. [10]
5. It is safest to assume that all people have a vicious streak and it will come out when they are given a chance.
6. Honesty is the best policy in all cases. [10]
7. There is no excuse for lying to someone else. [10]
8. It is hard to get ahead without cutting corners here and there.
9. All in all, it is better to be humble and honest than important and dishonest. [10]
10. When you ask someone to do something for you, it is best to give the real reasons for wanting it rather than giving reasons that carry more weight. [10]
11. Most people who get ahead in the world lead clean, moral lives. [10]
12. Anyone who completely trusts anyone else is asking for trouble.
13. The biggest difference between most criminals and other people is that criminals are stupid enough to get caught.
14. Most men are brave. [10]
15. It is wise to flatter important people.
16. It is possible to be good in all respects. [10]
17. Barnum was very wrong when he said that there's a sucker born every minute. [10]
18. Generally speaking, men won't work hard unless they're forced to do so.
19. People suffering from incurable diseases should have the choice of being put painlessly to death.
20. Most men forget more easily the death of their father than the loss of their property.

References

Allsopp, J. H., Eysenck, J., & Eysenck, S. B. (1991). Machiavellianism as a component in psychoticism and extraversion. *Personality and Individual Differences*, *12*, 29–41.

[10] Reverse coded.

Anastasi, A. (1990). *Psychological testing*. New York: McMillan.

Barkow, J., Cosmides, L., & Tooby, J. (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity and social history. *Games and Economic Behavior*, *10*, 122–142.

Bolton, G. (1991). A comparative model of bargaining: Theory and evidence. *American Economic Review*, *81*, 1096–1136.

Bolton, G. E., & Ockenfels, A. (2000). A theory of equity, reciprocity, and competition. *American Economic Review*, *90*, 166–193.

Boone, C., De Brabander, B., & van Witteloostuiijn, A. (1999). The impact of personality on behavior in five prisoner's dilemma games. *Journal of Economic Psychology*, *20*, 343–377.

Brown, D. (1991). *Human universals*. Philadelphia, PA: Temple.

Brown, W., & Moore, C. (2000). Is prospective altruist-detection an evolved solution to the adaptive problem of subtle cheating in cooperative ventures: Supportive evidence using the Wason selection task. *Evolution and Human Behavior*, *21*, 25–37.

Camerer, C., Johnson, E. J., Rymon, R., & Sen, S. (1993). Cognition and framing in sequential bargaining for gains and losses. In K. Binmore, A. Kirman, & T. Piero (Eds.), *Frontiers of game theory* (pp. 27–48). London: MIT Press.

Cherulnik, P. D., Way, J. H., Ames, S., & Hutto, D. B. (1981). Impressions of high and low Machiavellian men. *Journal of Personality*, *49*, 388–400.

Christie, R. (1970a). Scale construction. In R. Christie, & F. Geis (Eds.), *Studies in Machiavellianism* (pp. 10–33). New York: Academic Press.

Christie, R. (1970b). Social correlates of Machiavellianism. In R. Christie, & F. Geis (Eds.), *Studies in Machiavellianism* (pp. 314–336). New York: Academic Press.

Christie, R., & Geis, F. (1970). Implications and speculations. In R. Christie, & F. Geis (Eds.), *Studies in Machiavellianism* (pp. 339–358). New York: Academic Press.

Christie, R., Gergen, K. J., & Marlowe, D. (1970). The penny–dollar caper. In R. Christie, & F. Geis (Eds.), *Studies in Machiavellianism* (pp. 173–188). New York: Academic Press.

Clark, K., & Sefton, M. (2001). The sequential prisoner's dilemma: Evidence on reciprocation. *The Economic Journal*, *111*, 51–68.

Cleckley, H. M. (1955). *The mask of sanity: An attempt to clarify some issues about the so-called psychopathic personality*. St Louis, MO: Mosby.

Coleman, J. (1990). *Foundations of social theory*. New York: Vantage.

Cooper, S., & Peterson, C. (1981). Machiavellianism and spontaneous cheating in competition. *Journal of Research in Personality*, *14*, 70–75.

Cummins, D. D. (1999). Cheater detection is modified by social rank: The impact of dominance on the evolution of cognitive functions. *Evolution and Human Behavior*, *20*, 229–248.

Daderman, A. M. (1999). Differences between severely conduct-disordered juvenile males and normal juvenile males: The study of personality traits. *Personality and Individual Differences*, *26*, 827–845.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam.

Dunbar, R. (1996). *Grooming, gossip, and the evolution of language*. Cambridge, MA: Harvard University Press.

Effler, M. (1983). Konservatismus, Machiavellianismus: Validitäts- und Reliabilitätsuntersuchungen. *Zeitschrift fur Differentielle und Diagnostische Psychologie*, *4*, 79–85.

Epstein, S. (1979). The stability of behavior: I. On predicting most of the people most of the time. *Journal of Personality and Social Psychology*, *37*, 1097–1126.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning and experimental games with unique, mixed strategy equilibria. *American Economic Review*, *88*, 848–881.

Exline, R. V., Thibaut, J., Hickey, C. B., & Gumpert, P. (1970). Visual interaction in relation to Machiavellianism and an unethical act. In R. Christie, & F. Geis (Eds.), *Studies in Machiavellianism* (pp. 53–77). New York: Academic Press.

Eysenck, H., & Gudjonsson, G. H. (1989). *The causes and cures of criminality*. New York: Plenum Press.

Fehr, B., Samson, D., & Paulhus, D. R. (1992). The construct of Machiavellianism: Twenty years later. In C. D. Spielberger, & J. N. Butcher (Eds.), *Advances in personality assessment* (pp. 77–116). Hillsdale, NJ: Erlbaum.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*, 817–868.

Forsythe, R., Horowitz, J., Savin, E., & Sefton, M. (1994). Replicability, fairness and pay in experiments with simple bargaining games. *Games and Economic Behavior*, *6*, 347–369.

Frank, R. H. (1988). *Passions within reason*. New York: Morton.

Geis, F., & Christie, R. (1970). Overview of experimental research. In R. Christie, & F. Geis (Eds.), *Studies in Machiavellianism* (pp. 285–313). New York: Academic Press.

Geis, F., & Levy, M. (1970). The eye of the beholder. In R. Christie, & F. Geis (Eds.), *Studies in Machiavellianism* (pp. 210–234). New York: Academic Press.

Geis, F. L., & Moon, T. H. (1981). Machiavellianism and deception. *Journal of Personality and Social Psychology*, *41*, 766–775.

Gigerenzer, G. (1996). Rationality: Why social context matters. In P. B. Bates, & U. M. Staudinger (Eds.), *Interactive minds: Life-span perspectives on the social foundation of cognition* (pp. 319–346). Cambridge: Cambridge University Press.

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Why social context matters. *Cognition*, *43*, 127–171.

Grusec, J. E., Lockhart, R. S., & Walters, G. C. (1990). *Foundations of psychology*. Toronto: Copp Clark Pittman.

Gunnthorsdottir, A., Houser, D., McCabe, K., & Ameden, H. (2001). *Excluding free-riders improves reciprocity and promotes the provision of public goods*. University of Arizona: Working Paper.

Gupta, M. D. (1987). Machiavellianism of different occupational groups. *Indian Journal of Psychometry and Education*, *18*, 61–66.

Hare, R. D., & Craigen, D. (1974). Psychopathy and physiological activity in a mixed-motive game situation. *Psychophysiology*, *11*, 197–206.

Hare, R. D. (1993). *Without conscience: The disturbing world of the psychopaths among us*. New York: Guilford.

Harrell, W. A. (1980). Retaliatory aggression by high and low Machiavellians against remorseful and non-remorseful wrongdoers. *Social Behavior and Personality*, *8*, 217–220.

Harrell, W. A., & Hartnagel, T. (1976). The impact of Machiavellianism and the trustfulness of the victim on laboratory theft. *Sociometry*, *39*, 157–165.

Hayashi, N., Ostrom, E., Walker, J., & Yamagishi, T. (1999). Reciprocity, trust and the sense of control: A cross-societal study. *Rationality and Society*, *11*, 27–46.

Hirshleifer, J. (1987). On the emotions as guarantors of threats and promises. In J. Dupre (Ed.), *The latest on the best: Essays on evolution and optimality* (pp. 307–326). Cambridge, MA: MIT Press.

Hoffman, E., McCabe, K., & Smith, V. L. (1996a). Social distance and other-regarding behavior in dictator games. *American Economic Review*, *86*, 653–660.

Hoffman, E., McCabe, K., & Smith, V. (1996b). On expectations and the monetary stakes in ultimatum games. *International Journal of Game Theory*, *25*, 289–301.

Jones, G. E., & Kavanagh, M. J. (1996). An experimental examination of the effects of individual and situational factors on unethical behavioral intentions in the workplace. *Journal of Business Ethics*, *15*, 511–523.

Kaplan, H., & Hill, K. (1985). Food sharing among ache foragers. *Current Anthropology*, *26*, 233–245.

Kiyonari, T., Tanida, S., & Yamagishi, T. (2000). Social exchange and reciprocity: Confusion or a heuristic. *Evolution and Human Behavior*, *21*, 411–427.

Lykken, D. T. (1995). *The antisocial personalities*. Erlbaum, NJ: Hillsdale.

Marin, G. (1973). Behavioral interchange and the dogmatism and Machiavellianism scales. *Revista Latinoamericana de Psicologia*, *5*, 257–262.

McCabe, K., Rassenti, S., & Smith, V. L. (1996). Game theory and reciprocity in some extensive form bargaining games. *Proceedings of the National Academy of Science, November 1996*, 13421–13428.

McCabe, K., & Smith, V. L. (2000). A comparison of naïve and sophisticated subject behavior with game theoretic predictions. *Proceedings of the National Academy of Sciences*, *97*, 3777–3781.

McCabe, K., Smith, V. L., & LePore, M. (2000). Intentionality detection and "mindreading": Why does game form matter. *Proceedings of the National Academy of Science*, *97*, 4404–4409.

McHoskey, J. W., Worzel, W., & Szyarto, C. (1998). Machiavellianism and psychopathology. *Journal of Personality and Social Psychology*, *74*, 192–210.

Mealy, L. (1995). The sociobiology of sociopathy: An integrated evolutionary model. *Behavioral and Brain Sciences*, *18*, 523–599.

Meyer, H.-D. (1992). Norms and self-interest in ultimatum bargaining: The prince's prudence. *Journal of Economic Psychology*, *13*, 215–232.

Mischel, W. (1977). The interaction of person and situation. In D. Magnusson, & S. Endler (Eds.), *Personality at the crossroads: Current issues in interactive psychology* (pp. 333–352). Hillsdale, NJ: Erlbaum.

Mudrack, P. E., & Mason, E. S. (1995). More on the acceptability of workplace behaviors of a dubious ethical nature. *Psychological Reports*, *76*, 639–648.

Orbell, J. M., & Dawes, R. M. (1993). Social welfare, cooperators' advantage, and the option of not playing the game. *American Sociological Review*, *58*, 787–800.

Palfrey, T. P., & Prisbrey, J. E. (1997). Anomalous behavior in public goods experiments: How much and why. *American Economic Review*, *87*, 829–846.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, *83*, 1281–1302.

Rosenthal, R. (1990). How are we doing in soft psychology. *American Psychologist*, *45*, 775–777.

Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models on the intermediate term. *Games and Economic Behavior*, *8*, 164–212.

Sjöberg, L. (1997). Explaining risk perception: An empirical evaluation of cultural theory. *Risk Decision and Policy*, *2*, 113–130.

Slovic, P., & Peters, E. (1998). The importance of worldviews in risk perception. *Risk Decision and Policy*, *3*, 165–170.

Smith, V. L., & McCabe, K. (2001). Goodwill accounting in economic exchange. In G. Gigerenzer, & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 305–326). Boston, MA: MIT Press.

Tamborni, R., Stiff, J., & Zillmann, D. (1988). Preference for graphic horror featuring male versus female victimization: Personality and past film viewing experiences. *Human Communication Research*, *13*, 529–552.

Terhune, K. W. (1970). Personality in cooperation and conflict. In S. Paul (Ed.), *The structure of conflict* (pp. 193–229). New York: Academic Press.

Thaler, R. H. (1992). *The winner's curse: Paradoxes and anomalies of economic life*. New York: Free Press.

Tooby, J., & Cosmides, L. (1995). Mapping the evolved functional organization of mind and brain. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 1167–1178). Cambridge, MA: MIT Press.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–57.

Ueijo, C. K., & Wrightsman, L. S. (1967). Ethnic-group differences in the relationship of trusting attitudes and cooperative behavior. *Psychological Reports*, *20*, 563–571.

Vecchio, R. P., & Sussmann, M. (1991). Choice of influence tactics: Individual and orginazational determinants. *Journal of Organizational Behavior*, *12*, 73–80.

Wedekind, C., & Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, *288*, 850–852.

Wilson, D. S., Near, D., & Miller, R. R. (1996). Machiavellianism: A synthesis of the evolutionary and psychological literatures. *Psychological Bulletin*, *119*, 285–299.

Wilson, D. S., Near, D. C., & Miller, R. R. (1998). Individual differences in Machiavellianism as a mix of cooperative and exploitative strategies. *Evolution and Human Behavior*, *19*, 203–212.

Wrightsman, L. S., & Wuscher, M. L. (1974). *Assumptions about human nature: A social-psychological approach*. Monterey, CA: Brooks/Cole.

Wrightsman, L. S. (1966). Personality and attitude correlates of trusting and trustworthy behavior in a two-person game. *Journal of Personality and Social Psychology*, *4*, 328–332.

Znakov, V. V. (1995). Machiavellianism and the fib phenomenon. *Voprosy Psychologii*, *6*, 59–70.