Perugini, Marco; Tan, Jonathan H. W.; Zizzo, Daniel John

Working Paper
Which is the More Predictable Gender? Public Good Contribution and Personality

Discussion Paper, No. 236

Provided in Cooperation with:
European University Viadrina Frankfurt (Oder), Department of Business Administration and Economics

Suggested Citation: Perugini, Marco; Tan, Jonathan H. W.; Zizzo, Daniel John (2005) : Which is the More Predictable Gender? Public Good Contribution and Personality, Discussion Paper, No. 236, European University Viadrina, Department of Business Administration and Economics, Frankfurt (Oder)

This Version is available at:
http://hdl.handle.net/10419/23808

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
"Which is the More Predictable Gender?
Public Good Contribution
and Personality"

Marco Perugini
Jonathan Tan
Daniel John Zizzo
Which is the More Predictable Gender?  
Public Good Contribution and Personality*

Marco Perugini  
Department of Psychology, University of Essex, Colchester CO4 3SQ, United Kingdom  
mp@essex.ac.uk

Jonathan H.W. Tan  
Institute of Microeconomics, European University Viadrina, Frankfurt-Oder 15230, Germany  
tan@euv-ffo.de

Daniel John Zizzo  
School of Economics, University of East Anglia, Norwich NR4 7TJ, United Kingdom  
d.zizzo@uea.ac.uk

March 1, 2005

Abstract
Personality questionnaires have been used and can be used to predict behavior in economic settings. Using two sets of state-of-the-art measures from personality psychology (the Big Six) and social psychology (Social Value Orientation), we find that the behavior of men is predictable in the first half of a public good contribution experiment, whereas that of women is not. This result agrees with the reinterpretation of Carol Gilligan’s (1982) view that women are more sensitive to the context in which decisions are made.

JEL Classification Codes: J16, C91, H41.
Keywords: gender, context, personality, public goods.

* We thank Ramin Farinpour and Jan Heufer for the translations, Jan Heufer, Jessica Kaehler, and Claudia Vogel for the experimental assistance, and Michael Ashton and Kibeom Lee for the HEXACO-PI. The financial support of the Deutsche Forschungsgemeinschaft is gratefully acknowledged.
1. Introduction

There has been growing interest on whether and how women behave differently from men in economic settings (e.g., C. Bram Cadsby and Elizabeth Maynes, 1998; Catherine Eckel and Philip Grossman, 2001; James Andreoni and Lise Vesterlund, 2001). Theories of social preferences and of cooperation have dramatically different implications according to the mixture of types in the population, and it is natural to think of gender as a possible marker of inter-individual differences in behavior.\(^1\) Such differences are possibly attributable to explanations based on gender differences in personality. Economic researchers have, of late, considered the predictive value of questionnaires drawn from social psychology, such as Social Value Orientation (SVO, Theo Offerman, Joep Sonnemans and Arthur Schram, 1996), or (increasingly) from personality psychology, such as the Big Five personality factors (Bryan Caplan, 2003).

The main novelty of our paper is the finding of a new, and fairly striking, gender effect. While questionnaire measures have predictive power on the average public good contribution of men (at least in the first half of our experiment), women’s behavior cannot be predicted, or can be predicted much less well, on the basis of their questionnaire responses. We suggest that this may be due to a greater context sensitivity of women’s choices (see Carol Gilligan, 1982; Gilligan and Jane Attannucci, 1988; Cadsby and Maynes, 1998), and therefore the more tenuous nature of the link between questionnaires items and actual public good contribution responses.

Ours is also the first experimental study of public good contribution using a measure based on the Big Five, and the first that uses a personality framework and related measure that represents an advance over the Big Five, the Big Six (Kibom Lee and Michael Ashton, 2004), compared to what has been used so far in economic experiments. We use a standard SVO instrument as well, and so our results are not specific to the use of a single methodology.

Section 2 briefly provides some background. Sections 3 and 4 describe the experimental design and results, respectively. Section 5 provides a discussion and concludes.

2. Background

\(^1\) The title of one of the papers in this literature (James Andreoni and Lise Vesterlund, 2001: ‘Which is the Fair Sex? Gender Differences in Altruism’) has provided the inspiration for that of this paper.
A. Gender Effects on Contribution Gilligan (1982) claimed that women act more in terms of care and connectedness to other people, while men act more in terms of abstract justice, rights and obligations. Faced with a moral dilemma, an agent with a ‘care perspective’\(^2\) will consider it on its own merits, in a contextualized fashion that takes into account how the agent is related to the individuals who are involved in the dilemma. As a consequence, women may display different behavior in different contexts as a function of the contextualized features involved, whilst men will tend to display behavior that is less context-sensitive and more rule-based. As far as evidence was provided for Gilligan’s claim, it was based on open-ended interviews.

In public good contribution or similar settings, a common interpretation of Gilligan’s work has been that women should contribute more since they should care more about other people (Jean Stockard, Alphans Van De Kragt and Patricia Dodge, 1988; Jamie Brown-Kruse and David Hummels, 1993; Stephanie Seguino, Thomas Stevens and Mark Lutz, 1996). Empirical results have been, however, surprisingly ambiguous.\(^3\) Christophe Boone, Bert De Brabander and Arjen van Witteloostuijn (1999) and Kenneth Clarke and Martin Sefton (2001) found no statistically significant difference in average cooperation due to gender in Prisoner’s Dilemmmas. Kruse and Hummels (1993) found less contribution with women than with men, and argued that this is evidence against Gilligan’s hypothesis. Anatol Rapaport and Albert M. Chammah (1965) found that women and men contributed as much in the initial round of Prisoner’s Dilemmas, but with repetition women contributed less, particularly when playing with other women; in contrast, Clifford Nowell and Sarah Tinkler (1994) found more contribution in all-female groups in a linear public good contribution experiment. Results are ambiguous and often depend on the treatment (e.g., John Solow and Nicole Kirkwood, 2002, and Jean Stockard, Alphans van Kragt and Patricia Dodge, 1988). In a one-shot public good contribution setting, women contributed more than men (Seguino et al., 1996), which is broadly consistent with findings of greater initial cooperation by women that tend to disappear in a repeated setting (Charles Mason, Owen Phillips and Redington, 1991; Andreas

\(^2\) The terminology is from Gilligan and Attanucci (1988).

\(^3\) The ambiguity of the findings agrees with the complex picture also emerging from dictator and ultimatum games (Andreoni and Vesterlund, 2001; Eckel and Grossman, 2001; Avner Ben-Ner, Faminin Kong, Louis Putternam, 2004a; Ben-Ner, Putternam, Kong and Dan Magan, 2004b). For a more complete review than the one provided here see Eckel and Grossman (2001). For a review of psychological studies testing Gilligan’s hypothesis about moral reasoning, see Sara Jafee and Janet Shibley Hyde (2000).
Ortmann and Lisa Tichy, 1999; the linear public good contribution treatment of Cadsby and Maynes, 1998). There are a variety of factors that may explain some of the differences in the results, such as the greater risk aversion of women, the use of classroom subjects and more generally the social distance among participants and with the experimenter, the knowledge of the gender of the coplayers, and the potential of greater sensitivity by women to being let down.4

Does Gilligan’s theory really imply that we should expect more cooperation from women in experiments?5 The alternative interpretation by Cadsby and Maynes (1998) is that women are more sensitive and responsive to others’ play, and, as such, they are better able to coordinate around a selected equilibrium. They support this with evidence from a discrete threshold public good contribution experiment, and claim that the linear public good contribution setting should instead lead women to coordinate more quickly to the only available equilibrium which is one with complete free riding. What seems clear is that, in linear public good contribution games, if women do tend to cooperate more – other things being equal –, this may act as a brake to any tendency to coordinate to an equilibrium of complete non-cooperation. This leads to ambiguous results – and uninterpretable from the viewpoint of Gilligan’s theory -.

Greater context-sensitivity does seem to imply, however, that women should be more sensitive to social reinforcement in the form of deviations from the other players’ contributions. This prediction, also made by Cadsby and Maynes, will be verified in our experiment.

Context-sensitivity should also imply that, if women are asked personality questionnaires (only indirectly linked to a behavioral task they are faced with at some other point in time), a lower correlation between measures and behavior is to be expected. In other

---

4 Eckel and Grossman (2001) have an excellent discussion of some of these factors. For example, they note that cooperation is risky, and that women may reduce cooperation as a result in risky environments. This effect may moderate the propensity of women to behave more altruistically. For example, it may imply less cooperation in Prisoner’s Dilemmas, where only full cooperation (a very risky gamble) is possible in alternative to the safe choice of defecting.

5 One possible problem is that experimental groupings are short-lived affairs in relation to which the kind of care that women would supposedly be better at would not have the chance to develop: Solow and Kirkwood (2002) address this issue by considering natural groups (in the form of student societies), but, as they carefully acknowledge, their results are then hostage of self-selection bias issues, which may dominate other effects.
words, questionnaires should be less able to predict the behavior of women than they are of men. We test this hypothesis below.⁶

There has been very little work related to this hypothesis: ours is the first attempt to systematically test the hypothesis of predictability of behavior, and to do so in a strategic setting. On the basis of a post-experiment feedback questionnaire, Brown-Kruse and Hummels (1993) suggest that, although women ‘talk’ more altruistically, they do not behave as such. They have no formal test of this hypothesis (and, given the open-ended nature of the questionnaire, it would be difficult to test this).⁷ Kelly Brown and Laura Taylor (2000) found three times as large a difference between actual and stated difference in contribution to public goods for men than for women, suggesting that any discrepancy between personality measures and behavior cannot be attributed to a greater tendency of women to misrepresent in words what they want to do. In relation to dictator game data, Ben-Ner et al. (2004b) might be interpreted as implying (in their abstract) that women’s behavior is better explained by a Big Five questionnaire than that of men, but this does not appear to be the case if one looks more in detail at their analysis on this point.⁸

**B. Big Six and Social Value Orientation**

*Big Six.* The case for the relevance of personality dimensions for an analysis and prediction of economic behavior has been recently made by Caplan (2003)⁹. Caplan has convincingly argued that the developments in personality psychology allow economists to enrich their theories of preferences. In fact, personality dimensions can be considered as preferences, variable (heterogeneous) across individuals while stable within individuals. They can allow for more refined theoretical analyses as well as better predictions of economic choices and interactions. Caplan focused his attention especially on the so-called Big Five.

---

⁶ A more general way of interpreting our work is as a test of the consistency between what people claim to be (on paper) and how they actually behave (in the laboratory context). For an example of dissociation between verbal and behavioral responses, see Zizzo (2003).

⁷ Stockard et al. (1988) make a similar point, but again almost entirely informally.

⁸ Their two regressions in Table 3 reports coefficients and standard errors for each of the Big Five personality measure. Out of ten coefficient comparisons, at a glance there are five that are significantly different from each other. Two (their Extraversion coefficients) have relatively large coefficients for men but not for women; for one (Neuroticism in the ordered logit only), the reverse is true; for two (Conscientiousness), the coefficients are of the same magnitude. The reason why the regressions for women are statistically significant while those for men are not is to be plausibly attributed to the non-personality-related variables in the regressions (none of which are even close to significance in the regressions for men).

⁹ Similar considerations, albeit couched in a psychological language, have been made by Perugini and Boele de Raad (2001).
The Big Five provides a taxonomic (multi-dimensional) structure of stable human personality characteristics (this being one of the fundamental goals of personality psychology). Several independent studies in a range of countries and languages have repeatedly produced variants\(^\text{10}\) of the following personality factors (dimensions), namely (I) Extraversion, (II) Agreeableness, (III) Conscientiousness, (IV) Emotional stability (or Neuroticism), and (V) Intellect or Openness to Experience (for reviews, see De Raad, 2000, Lewis Goldberg, 1993, and Jerry Wiggins, 1996). However, these conclusions have recently been challenged (Ashton and Lee, 2001). New empirical evidence has consistently pointed out that six factors – the so-called Big Six – may offer a better taxonomy than the Big Five. These claims are based on recent studies and re-analyses of cross-cultural data (e.g., Ashton, Lee, Perugini, Piotr Szarota, Reinout De Vries, Lisa Di Blas, Kathleen Boies, and De Raad, 2004). In short, the Big Six appears to offer a more comprehensive and cross-culturally generalizable taxonomic structure of basic personality traits than the Big Five.

The differences between Big Six and Big Five include: i) the introduction of a new additional factor referring to the broad domain of morality and integrity that has been labeled Honesty/Humility, and ii) the identification of (slight) variants of Agreeableness and Emotional Stability. Another strength of the Big Six is that, by considering the sixth factor, the interpretation of the other five factors becomes cleaner\(^\text{11}\) than in the original five-factor framework (Ashton, Lee and Goldberg, 2004; Michael Ashton, Kibeom Lee, Reinout De Vries, Marco Perugini, Augusto Gnisci and Ida Sergi, 2004).

Lee and Ashton (2004) have developed an item questionnaire, the HEXACO-PI, specifically designed to measure the Big Six, in which each of the six factors is identified with four facets (more specific sub-dimensions), as listed in Appendix B. Of particular relevance in this context is the factor Agreeableness. This factor has been associated with motives and outcomes that are particularly relevant for economic behavior. Agreeable people are more highly motivated to maintain positive relations with other people (William Graziano, Lauri Jensen-Campbell and Elizabeth Hair, 1996) and adopt a prosocial and communal

---

\(^{10}\) Variation in the items, and in turn facets, that define certain dimensions give rise to such variants.

\(^{11}\) A “cleaner” interpretation of a factor is attained when its distinctive features become more coherently grouped into a qualitatively clear core. In other words, the essential feature of a factor becomes more evident and more differentiated from secondary features. An element of subjective theoretical evaluation in this judgment is unavoidable. However, there are some empirical consequences that can be observed. As we will argue later on, the results of this study for Agreeableness neatly illustrates the point.
orientation toward others rather than an antagonistic one (Oliver John and Sanjay Srivastava, 1999). Agreeableness is also associated with indicators of altruism (e.g., Ashton, Sampo Paunonen, Edward Helmes and Douglas Jackson, 1997) because, being motivated to maintain positive relations with others, agreeable persons are likely to help. These motives have been linked to a willingness to suspend one’s personal interest for the good of one’s social group (David Buss, 1991), and high Agreeableness should be associated with higher public good contribution (Van der Zee and Perugini, forthcoming).  

Note that there is a slight but relevant difference between Agreeableness as defined within the Big Five and as defined within the Big Six framework. In both frameworks, the key markers of Agreeableness are aspects such as gentleness, good-naturedness, agreeableness, tolerance, and patience. However, Agreeableness as defined in the Big Six includes content related to even temper, lack of irritability, and sensitivity, whereas these elements are part of the low Emotional Stability factor in the Big Five framework. In the Big Six framework, Agreeableness is defined in terms of four facets with self-explanatory labels: Forgiveness, Gentleness, Flexibility, and Patience.

**Social Value Orientation.** Departing from the more general basic personality framework offered by the Big Six, there is another important line of research that is especially relevant in this context. Several studies in social psychology have shown the relevance of considering one’s social value orientation or preference for particular patterns of distributions of outcomes for self and others (David Messick and Charles McClintock, 1968; Paul Van Lange, Wilma Otten, Ellen De Bruin and Jeffrey Joireman, 1997). A distinction has been made between individuals with (1) a prosocial (cooperator) value orientation (i.e., enhancing joint outcomes and enhancing equality in outcomes); (2) an individualistic orientation (i.e., enhancing outcomes for self with no regard to other’s outcome); and (3) a competitive orientation (i.e., enhancing relative advantage over other’s outcome). Whilst an individualistic orientation is

---

12 We may interpret this as a stronger positive weight placed on the payoffs of other individuals or the group as a whole, in the subjective utility function.

13 Technically speaking, the factorial space between Emotional Stability and Agreeableness in the Big Five is rotated of approximately 30 degrees in the Big Six. This rotation implies that some markers of the factors are different in the two frameworks. Another consequence of this rotation is that Emotional Stability is defined as Emotionality in the Big Six framework. This rotation allows a better understanding of the interplay of the different personality factors that are involved in prosocial behavior, as elegantly shown by Ashton and Lee (2001).

14 This latter orientation is usually relatively infrequent and therefore we will not devote much space to discuss it.
reflected in the standard assumptions of most economic models, it is important to highlight that cooperators construe social dilemmas, including public goods, differently. They tend to expect other people to cooperate and see social dilemmas as an occasion for cooperative exchanges. The relevance of this orientation for economic behavior has been already underscored by Offerman, Sonnemans and Schram (1996). In a step-level public good experiment, they found that cooperators (as defined by the SVO) systematically contribute more than individualists, albeit this difference tended to decrease in later contribution periods.

To sum up, two personality factors appear to be the best candidates to predict systematic differences in contribution in a public good environment. The first, Agreeableness, is one of the basic personality dimensions and features both in the Big Five and in the Big Six. Among the two, the cleanest definition of Agreeableness is provided by the Big Six framework. The second, Cooperators (as defined by the SVO), is a more specific dimension that is not part of the basic personality framework but that nonetheless appears ideally suited for predictions of economic behavior.

3. Experimental Design

The experiment took place in the experimental economics laboratory at the European University Viadrina, Frankfurt-Oder, Germany. A total of 108 subjects, with nine subjects in each session, participated in twelve sessions, six in November 2003 and six in November 2004. Subjects were recruited via email from a database of potential subjects, from the faculties of business administration and economics, cultural science, and law. We had 37 men and 71 women in our sample. None had prior experience in public good experiments. The interactive parts of the experiment were fully computerized. Standard experimental procedures were followed: computer terminals in the laboratory were individually partitioned, preventing subjects from seeing each other’s computer screens, or communicating by means of audio, facial, or body signals; neutral phrasings were used, such as ‘co-participant’ instead of ‘co-player’, ‘group’ instead of ‘team’; subjects could only participate to one session, and were duly paid at the end of it.

Due to the demographical composition of the university population, and the fact that we chose not to emphasize gender in the recruitment process (a point we shall return to below), we were not able to get a more balanced sample.
The experiment entailed two stages, each with 10 rounds of play. Subjects played in anonymously and randomly formed groups of three. Groups were re-matched for stage 2. Groups were composed of the same subjects within each stage. Group members were different across stages. In each round, subjects were allocated 20 ‘points’ (each worth €0.20) which they could either keep or invest. Each member in the group received 0.5 points for each point invested by any member of the group. Therefore, the design was one of a standard linear public good contribution game with ‘partners’ (Mark Isaac and James Walker, 1988), and, if we label $x_i$ is the contribution of player $i$, the payoff to player $i$ each round was equal to

$$20 - x_i + 0.5 \sum_{j=1}^{n} x_j,$$

Before stage 1 began, subjects completed instruction sheets and control questionnaires to check their understanding. They were allowed to begin only after the experimenters had checked the answers to these control questions; in the event of any incorrect answers, subjects were individually advised. At the top of the screen, the relevant stage, round, and participant number was displayed. At the end of each round, subjects received feedback on the investments and earnings of everyone in the group.

After completing stages 1 and 2, subjects answered a set of psychological questionnaires, inclusive of a Big Six and of a SVO questionnaire. At the end of the experiment, the computer informed subjects of their winnings, based on the points earned in two winning rounds – one from each stage 1 and 2 – randomly chosen by the computer. They were paid a fixed fee of €6 for completing the psychological questionnaires. There was no show-up fee. Subjects were informed of the winning rounds for each stage, and their payments after completing all the questionnaires. They were then paid in private. The mean payment was €15.04 for about 1.5 hours of work.

No emphasis was placed on gender in the recruitment process or while running the experiment. All our experimental sessions had both men and women, and subjects had no information about the gender of their coplayers.

---

16 We also had a questionnaire to probe reciprocity motivation (Perugini, Marcello Gallucci, Fabio Presaghi and Anna Paola Ercolani, 2003) and an alternative questionnaire probing for Big Five Agreeableness (the IPIP questionnaire, 2001). Neither was systematically correlated with behavior and they will not be referred further in this paper (except in footnotes 25 and 28 in relation to the IPIP).

17 We followed Andreoni and Vesterlund (2001) on this matter. As they note, the lack of emphasis on gender and the use of mixed groups may avoid additional psychological influences (such as male or female bonding) and
The experiment was conducted in German. English translations of the instructions are found in Appendix A, and of material relevant to the Big Six and SVO questionnaires in Appendix B.18

4. Experimental Results

A. Behavioral Patterns

Figure 1 plots mean contribution values for men and women participating in the experiment, and Table 1 contains the relevant data. The overall pattern of contributions is commonly observed in similar linear public good contribution games, displaying both a negative trend in contributions through time and a restart effect as subjects started stage 2 (Richard Cookson, 2000).

In stage 1 women contributed 5.97 tokens on average, about half a point more than men ($P < 0.1$ in a Mann-Whitney two-tailed test); in stage 2 men contributed slightly more, although not statistically significantly so in a Mann-Whitney test.

Table 2 reports the results of four of the regressions used to test for whether women converge to the equilibrium prediction of zero contribution more quickly. The dependent variables are, for models 1 and 2, the amount contributed, and, for models 3 and 4 (estimated with a logistic regression), a dummy variable equal to 1 if there was zero contribution and 0 otherwise. The independent variables are a gender dummy (equal to 1 for men), a round dummy (between 1 and 20), and, in models 2 and 4, the stage number (1 or 2). The stage variable is useful to control for the restart effect.

The picture from the regressions on average contribution is one where women contribute more than men, but less with time: controlling for the restart effect, the gender may reflect more the gender differences that would naturally occur in environments not explicitly designed to test for such differences. Of course, single-gender groups may be interesting in their own right.

18 The original questionnaires were in English. Careful precautions were taken in the translation procedure to maintain consistency. It was first translated to German (by Person 1), and then this was back-translated to English (by Person 2), to be compared with the original English version. Together with the authors, discrepancies were discussed and eliminated until consistency was achieved. Persons 1 and 2 are native speakers of German. Person 1 was an Economics undergraduate with study abroad experience in the USA. Person 2 was trained in English and Higher German throughout school, and holds Bachelors and Masters degrees in Law from a British university. Neither is an author for this paper.
dummy x interaction term is statistically significant. The evidence for greater convergence by women to zero contribution is on the borderline of significance.\textsuperscript{19}

Are women more sensitive to social rewards and punishments, or anyway “better” adaptive learners?\textsuperscript{20} The Spearman correlation between the change in contribution per round ($D_{Inv}$) and the difference from the mean contribution of the two other players in the previous round ($LDiff$) is equal to $-0.60$ for women ($P < 0.01$) but only to $-0.41$ for men ($P < 0.01$). Table 3 presents the results of two regressions testing for adaptive learning effects.\textsuperscript{21} Model 1 is simply a regression of the gender dummy, $LDiff$ and gender x $LDiff$ on $D_{Inv}$. Model 2 is a robustness test, which replaces $LDiff$ with $LDiff'$ (the difference between own contribution and the mean contribution of all three players in the previous round). It gives a similar picture. Figure 2 plots $D_{Inv}$ against $LDiff$ for men and women. While both men and women increase their contribution if they are socially rewarded, and decrease it if their contribution is punished, a difference in one point from the mean contribution of the other (or all) subjects, i.e. $LDiff$ (or $LDiff'$), leads to a mean contribution change, i.e. $D_{Inv}$, of -0.27 (or -0.40) points for men, but -0.38 (or -0.87) for women.

**B. Personality and Behavior**

*Social value orientation.* Using the method described in section 2, we can classify 47 subjects (44%) as cooperators, 43 subjects as individualists (40%) and 4 (4%) as competitive. Overall, some 87-88% of the subjects’ choices were consistent enough to be classifiable in one of these categories. The criterion for consistency is that at least 6 out of the 9 choices can be classified in the same category. The percentage of unclassified subjects is in line with the results found by Offerman et al. (1996) and Van Lange et al. (1997). It is noteworthy that, out of 14 subjects that were left unclassified, 13 were women and only 1 was a man. We shall return to this point below.

As in Offerman et al. (1996), given the small number of competitive subjects, our focus is on the comparisons between cooperators and individualists; we have less individualists,

\textsuperscript{19} In model 3 (model 4) the coefficient on gender x round is significant at $P = 0.05$ ($P = 0.06$) if a one-tailed test based on Cadsby and Maynes’ predictions is used.

\textsuperscript{20} Since payoffs map onto actions, and one’s payoff is lower (higher) relative to the other’s/group’s average if one’s contribution is higher (lower), a “better adaptive learner” is by definition one who adjusts contributions in the opposite direction of contributions relative to the other’s/group’s average contributions in the previous period.

\textsuperscript{21} For obvious reasons, such as the large action set and the small number of repetitions, we restrict our attention to a learning model with memory lag of one period.
however, than the 65% of Offerman et al.’s sample. In the overall sample cooperators contribute 6.41 points on average against the 3.71 of competitors ($P < 0.001$, Mann-Whitney test), seemingly confirming Offerman et al.’s finding that SVO can be used to predict behavior in public good contribution settings.

However, if we decompose the data by gender, we find that the predictive power is largely driven by the sample of men (see Figure 3 and Table 4). 22 Men cooperators contributed a mean 7.30 points, over twice as much as men individualists (2.99, $P < 0.01$ in a Mann-Whitney test); conversely, women cooperators contributed 5.95 points, less than two points more than women individualists (4.22, $P < 0.05$ in a Mann-Whitney test).

Another way of looking at the predictive power of SVO measures, which avoids discarding any observations from our sample, is to verify whether cooperators’ contributions, or individualists’ contributions, are different from those of everyone else pooled together. From Figure 3, this is clearly the case for men ($P < 0.005$ in Mann-Whitney tests) but less clearly so for women (only $P < 0.1$ in Mann-Whitney tests).

Table 4 also reports mean contribution values in stage 1 and stage 2 of the experiment. There is less predictive power in stage 2 than in stage 1, but men cooperators still appear to contribute significantly more than men individualists, and more than everyone else ($P < 0.1$ or better in Mann-Whitney tests). Conversely, in stage 2 SVO does not have statistically significant predictive power for women.

In conclusion, the SVO classification method allows to pick up the behavior of men much better than that of women: virtually all the subjects whose SVO choices did not lend themselves into classification in a preference type were women, and SVO appears a much better predictor of public good contribution choices with men than with women.

**Big Six.** Table 5 summarizes mean values for each of the Big Six dimensions and, in more detail, for the Agreeableness facets. 23 To what extent is it possible to use the Big Six to predict public good contribution for men and for women?

To answer this question, Table 6 reports Spearman correlation coefficients between mean contribution values and the Big Six dimensions and Agreeableness facets. 24 A first

---

22 31 women were classified as cooperators and 16 as individualists; 25 men were classified as cooperators and 18 as individualists.
23 Two-tailed Mann-Whitney tests find statistically significant differences in mean reported values in relation to Honesty ($P < 0.05$), Emotionality ($P < 0.001$), Extraversion ($P < 0.06$) and Forgiveness ($P < 0.05$): women report themselves as more honest, more emotional, less extroverted and less forgiving than men.
finding is that, insofar as there is predictive power of personality variables, this is largely
driven by stage 1. A second finding is that, in aggregate and largely in relation to stage 1,
more honest and agreeable subjects are likely to contribute more to the public good (see
Section 2); of the four facets of Agreeableness, Flexibility and Patience have the greatest
predictive power, but all four facets are significant at \( P < 0.07 \) or better in two-tailed tests or \( P
< 0.05 \) in one-tailed tests.\(^{25}\) Spearman \( \rho \) coefficients, where significant, hover around the 0.2-
0.3 range.

As we decompose the analysis by gender we move to our third finding from Table 6:
the moderate size aggregate correlation coefficients hide a stark difference between the
predictive power of personality measures for men relative to women. Personality measures
can be powerful predictors of the public good contribution of men in the first stage, but they
are mostly useless as predictors of behavior for women. More forgiving, flexible and patient
men are likely to engage in more public good contribution in stage 1, with Spearman \( \rho 
\) coefficients equal to 0.525 \( (P = 0.001) \), 0.368 and 0.359 \( (P < 0.05) \), respectively; the overall
Agreeableness score has Spearman \( \rho = 0.496 \) \( (P < 0.01) \) for stage 1. Figure 4 depicts these
relationships and contrasts them with the much weaker ones found for women. Table 7
reports the results of four regressions that we did to test formally whether the predictive
power of the above mentioned personality dimensions was significantly different for men and
women in predicting the mean contribution in Stage 1. Each regression included the dummy
gender (1 for males), a personality factor, and the interaction term gender by personality.\(^{26}\) A
significant interaction term implies a statistically different relation between the personality
factor and the mean contribution for men and women. Three of the four interaction terms are
statistically significant at \( P < 0.05 \) (Agreeableness, Forgiveness, Patience), whereas the fourth

\(^{24}\) For the sake of comparability, we also computed Spearman correlation coefficients between SVO variables
and mean contribution. Unsurprisingly, the results are exactly along the same lines as those that can be found
using Mann-Whitney tests.

\(^{25}\) In congruence to the evidence for the superiority of the Big Six over the Big Five classification (see section
2.B), the Big Six Agreeableness works better than the Big Five Agreeableness in predicting behavior. Spearman
correlation coefficients with mean contribution are 0.123, 0.173 and – 0.009 for the whole sample, stage 1 and
stage 2 respectively (in relation to stage 1, \( P < 0.1 \)). The Big Five Agreeableness score can be derived from
the Big Six data as: \( 2 \times (\text{Honesty-Humility}) + 2 \times (\text{Agreeableness}) + \text{Dependence} + \text{Sentimentality} + \text{Sociability}
+ \text{Liveliness} \) (see Kibeom Lee and Michael Ashton, 2005). The IPIP measure of Agreeableness (see footnote 16) is
a standard measure for the factor in the Big Five classification and in our version contained 78 items measuring
10 facets. The results cannot be explained therefore because we used a less reliable or thorough measure of the
Big Five Agreeableness.

\(^{26}\) The personality variables have been centered prior to computing the interaction term, so as to eliminate
undesirable multicollinearity effects (e.g., Donald Marquardt, 1980).
one is marginally significant (P < 0.07, Flexibility). A simple effect analysis, as depicted in Figure 4, shows that the effects are both consistent and sizeable. For instance, considering Agreeableness, a change of one unit in the score for this trait (range 1-5) implies a change in tokens invested in the public good of 5.12 for a male and only 1.23 for a female, therefore with a ratio that is more than 4 times greater for males.

To summarize, we find that, particularly in relation to stage 1, personality dimensions based on Agreeableness have predictive power on the average contribution of men. The same is not true for women.

5. Discussion and Conclusions

The aggregate behavioral findings are not particularly surprising, as they only show a slightly higher cooperation rate for women relative to men, especially in first half of the experiment. Figure 1 also showed some negative trend in contributions, not unlike that found for example in Mason et al. (1991).

Our key finding was that, using personality questionnaires, we could predict the behavior of men to some extent, especially in the first half of the experiment, but not that of women. Another dimension in which personality questionnaires seemed to work less well was that, with only one exception, all of the subjects whom we were not able to classify as belonging to a SVO personality type were women. Both findings suggest that personality questionnaires are less useful for economic decision-making with women than with men.

A possible criticism may be that our study shows a more general inadequacy of personality measures to predict behavior. However, this criticism would neglect a considerable body of evidence that supports the empirical validity and usefulness of considering personality dimensions. For instance, hundreds of studies analyzed with meta-analytical statistical techniques have provided robust evidence that personality traits play an important role in predicting effective leadership (Joyce Bono and Timothy Judge, 2004), job performance (Joyce Hogan and Brent Holland, 2003; Timothy Judge and Ilies Remus, 2002) and job satisfaction (Timothy Judge, Daniel Heller and Michael Mount, 2002). Within the economic literature, the analysis of the role played by personality traits is still in its infancy, but there are examples that might pave the way for a broader application of a personality analysis. For instance, besides the already mentioned contribution of Caplan (2003), where he
advocates the adoption of personality traits to model heterogenous preferences, Samuel Bowles, Herbert Gintis and Melissa Osborne (2001) have argued for considering some personality traits as the basis of incentive enhancing preferences that can be usefully applied to model some specific principal-agent relationships. Personality traits have also started to be considered in experimental economics. For instance, besides the already mentioned work of Offerman et al. (1996) with the SVO, the influence of personality traits has been investigated by Robert Kurzban and Daniel Houser (2001) in a circular public good setting, by Susanne Büchner, Giorgio Coricelli and Ben Greiner (forthcoming) in a solidarity game, by Hermann Brandstätter and Manfred Königstein (2001) in an ultimatum game with advanced production, and by Ben-Ner et al. (2004b) in a dictator game.

It is, nevertheless, fair to note that the results have not always shown supporting evidence. It is an open question why our results differ from those in Ben-Ner et al. (2004b), where the Big Five do not predict the behavior of men any better than that of women, and less variance appears to be explained by the Big Five anyway. We believe that our results may be less in conflict than it may seem at first glance, since they would also have been less strong had we used the Big Five, as Ben-Ner et al. did, rather than the Big Six. Our conjecture is that, had Ben-Ner et al. used the Big Six and so a cleaner measure of Agreeableness and its facets, they would have found stronger results, as we do.

A second objection is that economists need not be interested in personality as obtained through questionnaire measures but simply in behavior. This objection fails to recognize that it is exactly behavior we are interested in, and, given the heterogeneity of behavioral attitudes in important settings such as public good contribution or bargaining games, it makes sense to use questionnaires to measure personality dimensions that can help to predict behavior. Gender is an obvious marker for possible differences in behavior, and one of considerable importance in the light of economic applications such as labor market differences in wages and career prospects (Donna Ginther and Kathy Hayes, 2003), where “personal

27 Big Five Agreeableness performs less well on our data than Big Six Agreeableness: see footnote 25.
28 As already mentioned in footnote 25, our experimental dataset adds to a growing body of evidence showing that the Big Six outperforms the Big Five. An alternative or complementary explanation for the discrepancy between our results and Ben-Ner et al. is that their Big Five questionnaire (the NEO-FFI) contains only sixty items to measure all five factors. To put this in perspective, our Big Six questionnaire had 192 items and our measure of Agreeableness alone contained 32 items. This line of explanation of the discrepancy cannot, however, explain why one of the other questionnaires which we used (IPIP Agreeableness, as mentioned in footnote 16) also performed poorly, even though it had 78 items to measure Big Five Agreeableness.
characteristics” such as cultural background, and tastes for home and work also play a role (Heather Antecol, 2001). Of course, alternative methods to capture heterogeneity of preferences exist and should be used. We do not see this issue as an ‘either-or’ choice. Questionnaire-based personality measure should be considered as a complement and not a substitute of other behavioral measures. For instance, in our study the role played by personality seems largely restricted to the first half of the experiment, before the feedback and learning from the experiment acquires a greater weight in determining choices. Nonetheless, the interplay between personality traits and learning dynamics appears as a very interesting avenue for future research.

We can make sense of the lesser validity of questionnaire measures in the case of women in two ways: by stating that women are more likely to misrepresent their preferences, or by stating that women are more likely to be context-sensitive in their preferences. In the first case the issue would be one of preference revelation, in the second one of true preferences. There are three reasons that make us incline to believe in the latter rather than the former. First, contingent valuation experiments have failed to find any evidence of greater misrepresentation of preferences by women, and, if anything, the study by Brown and Taylor (2000) shows exactly the reverse. Second, the greater context-sensitivity of preferences fits with the evidence of Cadsby and Maynes (1998) and with the finding of greater sensitivity of women to social rewards and punishments by the other players, which we find in our data. Third, our findings are consistent with the results of Kurzban and Houser (2001). The authors classified the participants in the game in four types depending on their pattern of choices in the circular public good game. Interestingly, more men than women were classified as cooperators (who contribute 100%) whereas more women than men were classified as conditional cooperators. More specifically, the adjusted probability of being a cooperator was 0.41 for men and 0.19 for women, whereas of being a conditional cooperator was 0.29 for men and 0.49 for women. In other words, women were much more likely to be sensitive to the other players’ choices and cooperate according to the emerging aggregated behavior. But, of course, being more sensitive to other players’ choices implies that one is relatively less prone to follow one’s own personal inclination – as emerging from the responses to questionnaire measuring basic personality dimensions. The empirical consequence is that for women these

29 For an important recent attempt, see Roberto Burlando and Francesco Guala (2002).
personality measures will be less predictive of behavioral choices. In short, these reasons converge with our data in supporting an interpretation of women’s behavior in terms of greater contextual sensitivity. “La donna e’ mobile” [a woman is fickle], sings Rigoletto in Verdi’s “La traviata”. Provided it is understood in its correct positive meaning as greater context sensitivity and not as a pejorative stereotype, this sentence captures succinctly the main message of this paper.
Appendix A. Experimental Instructions

Stage 1 Instructions

You are about to participate in an experiment on decision-making. The experiment is divided into 2 stages, and a set of questionnaires has to be filled after the two stages. You will be paid your earnings from two winning rounds randomly chosen by the computer, one from each stage (1 and 2), plus 6 euros for the set of questionnaires. You will not know which rounds are winning rounds until the end of the experiment.

Stage 1 and 2 consists of 10 rounds of the same task. In each stage, the computer will randomly match you and the other experimental participants into 3 groups of 3 participants each, so there will be 2 co-participants in your group for each stage. The other 2 co-participants in your group stay the same from round to round during the stage. However, co-participants will change across stages: put it differently, you will be a member of a different group in Stage 2 relative to Stage 1. In Stage 2, you will not be matched again with any co-participant that you have been matched with in Stage 1.

Task: You are given 20 points. You may decide to invest any amount from 0 to 20 points; in addition, you will keep whatever is not invested. For each point that you and the co-participants from your group invest, every one of you will receive 0.5 points for the total amount invested. In general, you will receive:

\[ \text{Payment} = 20 - \text{Your Investment} + 0.5 \times \text{Total Investment in your Group} \]

Example: If you invested 9 points, co-participant A invested 11 points, and co-participant B invested 10 points, the total amount invested is 30 points. For this round, you will have 26 points, co-participant A will have 24 points, and co-participant B will have 25 points as the final earnings for the round.

Once you have decided how much to invest, please type in the number (from 0 to 20) into the text box provided. To confirm your decision, click on Confirm, where a message box will appear asking if you are sure of your decision. If so, click on OK in the message box, and then on Confirm again. If not, you may change your decision either by double clicking on the text box, or clicking on Cancel, and then entering your new decision. You may change your decision anytime before the 2nd click on Confirm.

After all your co-participants have made their decisions, you will be told how much each of you invested, and how much each of you earned. Each point that you have earned is worth 0.20 euros. So for the above example, if it is chosen as a winning round, you will earn 5.20 euros for this round. You may then move on to the next round by clicking on Continue.

At the end of the stage, you will be prompted by the computer to call for the experimenter. After reading these instructions, please answer the questions printed on the next page, and call for the experimenter before you begin. Many thanks and good luck!
Appendix B. Questionnaires.

a) Social Value orientation (SVO)

In this part of the booklet we ask you to make some choices. You can do this by circling the letters A, B, or C. Your choices will determine the number of points that you and someone else will receive. Imagine that this other person also makes choices in exactly the same task.

Who is this other person?
You must try to imagine that this other is a person who you do not know (who you have never met) and that you will never meet (or get to know) this person. So this other person is completely unknown to you.

What is the meaning of these points?
The points represent something which is valuable to you: the more you accumulate, the better. The same applies to the other person: the more he or she accumulates, the better for him/her.

An example of this decision task is displayed below.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>you get</td>
<td>500</td>
<td>500</td>
<td>550</td>
</tr>
<tr>
<td>other gets</td>
<td>100</td>
<td>500</td>
<td>300</td>
</tr>
</tbody>
</table>

In this example, if you choose A, you would receive 500 points and the other would receive 100 points; if you choose B, you would receive 500 points and the other would receive 500 points; if you choose C, you would receive 550 points and the other would receive 300 points. We will ask you to make nine decisions.

Keep the following in mind:

- There are no right or wrong answers. You choose to circle A, B, or C, depending on which alternative you find most attractive, for whatever reason.

- Imagine that the points are valuable to you: the more you accumulate, the better. The same applies to the other person: the more points he or she accumulates, the better for him/her.
For each of the following nine choice situations, circle A, B, or C, depending on which column you prefer most.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) YOU</td>
<td>480</td>
<td>540</td>
<td>480</td>
</tr>
<tr>
<td>Other</td>
<td>80</td>
<td>280</td>
<td>480</td>
</tr>
<tr>
<td>(2) YOU</td>
<td>560</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>Other</td>
<td>300</td>
<td>500</td>
<td>100</td>
</tr>
<tr>
<td>(3) YOU</td>
<td>520</td>
<td>520</td>
<td>580</td>
</tr>
<tr>
<td>Other</td>
<td>520</td>
<td>120</td>
<td>320</td>
</tr>
<tr>
<td>(4) YOU</td>
<td>500</td>
<td>560</td>
<td>490</td>
</tr>
<tr>
<td>Other</td>
<td>100</td>
<td>300</td>
<td>490</td>
</tr>
<tr>
<td>(5) YOU</td>
<td>560</td>
<td>500</td>
<td>490</td>
</tr>
<tr>
<td>Other</td>
<td>300</td>
<td>500</td>
<td>90</td>
</tr>
<tr>
<td>(6) YOU</td>
<td>500</td>
<td>500</td>
<td>570</td>
</tr>
<tr>
<td>Other</td>
<td>500</td>
<td>100</td>
<td>300</td>
</tr>
<tr>
<td>(7) YOU</td>
<td>510</td>
<td>560</td>
<td>510</td>
</tr>
<tr>
<td>Other</td>
<td>510</td>
<td>300</td>
<td>110</td>
</tr>
<tr>
<td>(8) YOU</td>
<td>550</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>Other</td>
<td>300</td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td>(9) YOU</td>
<td>480</td>
<td>490</td>
<td>540</td>
</tr>
<tr>
<td>Other</td>
<td>100</td>
<td>490</td>
<td>300</td>
</tr>
</tbody>
</table>
HEXACO-PI\textsuperscript{30}

The following list contains descriptions that you can consider more or less true for yourself and with which you can agree or disagree. There are no right or wrong answers. We just ask you to answer how much you agree or disagree with each statement. The answer choices are:

1. STRONGLY DISAGREE
2. DISAGREE
3. NEUTRAL
4. AGREE
5. STRONGLY AGREE

We would like to remind you:

a) It is very important that you answer to every description. Please, check at the end whether all questions are answered.

b) We know that people may behave in a given way in a situation and in another way in a different situation. When deciding whether the description applies to you, you may think “It depends”. We therefore ask you to decide whether usually or on average the description applies to you.

Don’t take too much time to think about your answers.

\textit{Reported below are one example of items for each facet of each construct. Items preceded by a minus sign between brackets are reverse coded. For a definition of the facets see Lee and Ashton (2004, Table 1). For the full HEXACO-PI questionnaire, please contact Mike Ashton (mashton@brocku.ca) or Kibeom Lee (kibeom@ucalgary.ca).}

\textbf{Honesty-Humility}

\textbf{Sincerity:}
(-) If I want something from a person I dislike, I will act very nicely toward that person in order to get it.

\textbf{Fairness:}
I wouldn't cheat a person even if he or she was a real "sucker".

\textbf{Greed Avoidance:}
Having a lot of money is not especially important to me.

\textbf{Modesty:}
(-) Sometimes I feel that laws should not apply to someone like me.

\textbf{Emotionality}

\textbf{Fearfulness:}
(-) Where physical pain is involved, I’m a very tough person.

\textbf{Anxiety:}
Sometimes I feel nervous without really knowing why.

\textbf{Dependence:}
(-) I can "tough it out" on my own through any kind of personal hardship.

\textbf{Sentimentality:}

\textsuperscript{30} Copyright Kibeom Lee and Michael C. Ashton; printed with permission.
When someone close to me is concerned about something, I feel concerned too.

**Extraversion**

**Expressiveness:**
My style of speaking is often quite dramatic.

**Social Boldness:**
(-) I rarely express my opinions in group meetings.

**Sociability:**
I prefer jobs that involve active social interaction to those that involve working alone.

**Liveliness:**
(-) People often tell me that I should try to cheer up.

**Agreeableness**

**Forgiveness:**
(-) If someone who has been unkind to me starts being nice, I remain suspicious of that person for a long time.

**Gentleness:**
I tend to be lenient in judging other people.

**Flexibility:**
I often cooperate with other people even when I think their plans don't make sense.

**Patience:**
(-) I react very angrily if I find that someone is trying to cheat me.

**Conscientiousness**

**Organization:**
(-) People often joke with me about the messiness of my room or desk.

**Diligence:**
People sometimes call me a "workaholic".

**Perfectionism:**
(-) I don't like to spend time perfecting work that is already good enough.

**Prudence:**
I usually stop myself before doing anything that I might later regret.

**Openness to Experience**

**Aesthetic Appreciation:**
Sometimes I like to just watch the wind as it blows through the trees.

**Inquisitiveness:**
(-) I find TV nature programs to be very boring.

**Creativity:**
(-) I prefer doing things the way I've always done them, rather than waste time looking for a new way.

**Unconventionality:**
I like hearing about opinions that are very different from those of most people.

**Bibliography**


Lee, Kibeom, and Ashton, Michael C. *HEXACO Personality Inventory manual*. 2005, unpublished manuscript.


Figure 1. Average Contribution by Men and Women

![Graph showing average contribution by men and women over rounds.](image)

Figure 2. Change in Contribution Against Deviation from Other Players’ Mean Contribution

![Scatter plots showing change in contribution against deviation.](image)

DInv = change in contribution relative to previous round. LDiff = difference of own contribution from the mean contribution of the two other players in the previous round.
Figure 3. Mean Contribution Levels by Cooperative and Individualist Types

Sample sizes: cooperative women $n = 31$; individualist women $n = 25$; cooperative men $n = 16$; individualist men $n = 18$; others $n = 18$. 
Figure 4. Predictability of Stage 1 Mean Contributions Using Agreeableness Measures
Table 1. Evolution of contributions over time

| Round | Women | | | Women | | | Men | | | Men |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|       | Stage 1 | Stage 2 |       | Stage 1 | Stage 2 |       |       |       |       |       |       |
| 1     | 8.18   | 6.93   |       | 8.51   | 5.97   |       |       |       |       |       |       |
| 2     | 7.68   | 5.61   |       | 7.24   | 5.76   |       |       |       |       |       |       |
| 3     | 7.28   | 5.06   |       | 5.92   | 4.84   |       |       |       |       |       |       |
| 4     | 6.37   | 4.56   |       | 4.59   | 5.59   |       |       |       |       |       |       |
| 5     | 5.9    | 4.62   |       | 5.86   | 5.97   |       |       |       |       |       |       |
| 6     | 6.1    | 3.76   |       | 4.68   | 5.46   |       |       |       |       |       |       |
| 7     | 5.28   | 3.89   |       | 4.84   | 4.35   |       |       |       |       |       |       |
| 8     | 4.55   | 3.9    |       | 5.24   | 4.35   |       |       |       |       |       |       |
| 9     | 4.61   | 2.7    |       | 4.35   | 4.16   |       |       |       |       |       |       |
| 10    | 3.8    | 2.2    |       | 3.24   | 3.05   |       |       |       |       |       |       |
| Mean  | 5.97   | 4.32   |       | 5.45   | 4.95   |       |       |       |       |       |       |

Table 2. Regressions on Zero Contribution (Models 1 and 2) and Probability of Zero Contribution (Models 3 and 4)

| Contribution in Round | | | Probability of Zero Contribution | | | |
|-----------------------|-----------------|----------------|-----------------|-----------------|----------------|
| Model 1               | Model 2         | Model 3        | Model 4         |
| Constant              | 7.63            | -3.1           | -2.42           | 0               |
| Gender                | -1.13           | 1.13           | 1.11            | 0               |
| Round                 | -0.24           | 0.15           | 0.22            | 0               |
| Gender x Round        | 0.11            | -0.03          | -0.03           | 0.1             |
| Mean                  | 5.56            | Constant       | 2.92            | 0.99            |

Sample size: $n = 2160$. The $R^2$ for models 1 and 2 is 0.048 and 0.065 respectively. The $R^2$ for models 3 and 4 is 0.164 and 0.177 respectively. All $p$ values provided are two-tailed for the sake of comparability.
Table 3. Regressions on the Change in Contribution by Women Across Rounds

<table>
<thead>
<tr>
<th>Model 1</th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.55</td>
<td>-5.19</td>
<td>0</td>
</tr>
<tr>
<td>Gender</td>
<td>0.14</td>
<td>0.79</td>
<td>0.43</td>
</tr>
<tr>
<td>LDiff</td>
<td>-0.58</td>
<td>-25.52</td>
<td>0</td>
</tr>
<tr>
<td>Gender x LDiff</td>
<td>0.31</td>
<td>9.01</td>
<td>0</td>
</tr>
</tbody>
</table>

Model 2

<table>
<thead>
<tr>
<th>Model 2</th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.55</td>
<td>-5.19</td>
<td>0</td>
</tr>
<tr>
<td>Gender</td>
<td>0.14</td>
<td>0.79</td>
<td>0.43</td>
</tr>
<tr>
<td>LDiff</td>
<td>-0.87</td>
<td>-25.52</td>
<td>0</td>
</tr>
<tr>
<td>Gender x LDiff</td>
<td>0.47</td>
<td>9.01</td>
<td>0</td>
</tr>
</tbody>
</table>

Sample size: n = 2160. The $R^2$ for both models 1 and 2 is 0.28. LDiff = difference of own contribution from the mean contribution of the two other players in the previous round. LDiff' replaces mean contribution of the other players with the mean contribution of all three players in the group.

Table 4. Mean Contribution Levels

<table>
<thead>
<tr>
<th>Mean Public Good Contribution</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Whole Sample</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperative Women</td>
<td>6.7</td>
<td>5.2</td>
<td>5.95</td>
<td>31</td>
</tr>
<tr>
<td>Individualist Women</td>
<td>5.07</td>
<td>3.38</td>
<td>4.22</td>
<td>25</td>
</tr>
<tr>
<td>Other Women</td>
<td>5.98</td>
<td>4.09</td>
<td>5.03</td>
<td>15</td>
</tr>
<tr>
<td>Cooperative Men</td>
<td>7.77</td>
<td>6.83</td>
<td>7.3</td>
<td>16</td>
</tr>
<tr>
<td>Individualist Men</td>
<td>3.16</td>
<td>2.82</td>
<td>2.99</td>
<td>18</td>
</tr>
<tr>
<td>Other Men</td>
<td>6.83</td>
<td>7.7</td>
<td>7.27</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 5. Mean Big Six and Agreeableness Facets Values

<table>
<thead>
<tr>
<th>Agreeableness</th>
<th>Women</th>
<th>Men</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agreableness</td>
<td>2.78</td>
<td>2.84</td>
<td>2.8</td>
</tr>
<tr>
<td>Honesty</td>
<td>3.56</td>
<td>3.26</td>
<td>3.45</td>
</tr>
<tr>
<td>Emotionality</td>
<td>3.4</td>
<td>3</td>
<td>3.26</td>
</tr>
<tr>
<td>Extraversion</td>
<td>3.28</td>
<td>3.48</td>
<td>3.35</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>3.4</td>
<td>3.41</td>
<td>3.4</td>
</tr>
<tr>
<td>Openness to Experience</td>
<td>3.39</td>
<td>3.41</td>
<td>3.4</td>
</tr>
<tr>
<td>Forgiveness</td>
<td>2.32</td>
<td>2.61</td>
<td>2.42</td>
</tr>
<tr>
<td>Gentleness</td>
<td>2.95</td>
<td>2.91</td>
<td>2.94</td>
</tr>
<tr>
<td>Flexibility</td>
<td>2.8</td>
<td>2.82</td>
<td>2.81</td>
</tr>
<tr>
<td>Patience</td>
<td>3.03</td>
<td>3.04</td>
<td>3.03</td>
</tr>
</tbody>
</table>

Forgiveness, Gentleness, Flexibility and Patience are facets (aspects) of Agreeableness.
Table 6. Nonparametric Correlations between Personality Measures and Mean Contribution

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th></th>
<th></th>
<th>Men</th>
<th></th>
<th></th>
<th>Women</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Stage 1</td>
<td>Stage 2</td>
<td>Total</td>
<td>Stage 1</td>
<td>Stage 2</td>
<td>Total</td>
<td>Stage 1</td>
<td>Stage 2</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>0.252***</td>
<td>0.289***</td>
<td>0.088</td>
<td>0.402**</td>
<td>0.496***</td>
<td>0.087</td>
<td>0.183</td>
<td>0.218*</td>
<td>0.093</td>
</tr>
<tr>
<td>Honesty</td>
<td>0.134</td>
<td>0.191**</td>
<td>0.034</td>
<td>0.188</td>
<td>0.249</td>
<td>-0.034</td>
<td>0.02</td>
<td>0.07</td>
<td>0.022</td>
</tr>
<tr>
<td>Emotionality</td>
<td>-0.027</td>
<td>-0.004</td>
<td>-0.022</td>
<td>-0.057</td>
<td>-0.113</td>
<td>0.013</td>
<td>-0.135</td>
<td>-0.058</td>
<td>-0.132</td>
</tr>
<tr>
<td>Extraversion</td>
<td>-0.151</td>
<td>-0.125</td>
<td>-0.15</td>
<td>-0.159</td>
<td>-0.081</td>
<td>-0.218</td>
<td>-0.067</td>
<td>-0.121</td>
<td>-0.055</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>-0.002</td>
<td>-0.045</td>
<td>0.02</td>
<td>-0.188</td>
<td>-0.202</td>
<td>-0.162</td>
<td>0.155</td>
<td>0.093</td>
<td>0.142</td>
</tr>
<tr>
<td>Openness to Experience</td>
<td>0.27</td>
<td>0.054</td>
<td>0.054</td>
<td>0.207</td>
<td>0.206</td>
<td>0.167</td>
<td>-0.033</td>
<td>-0.012</td>
<td>-0.009</td>
</tr>
<tr>
<td>Forgiveness</td>
<td>0.156</td>
<td>0.201**</td>
<td>0.053</td>
<td>0.287*</td>
<td>0.525***</td>
<td>0.009</td>
<td>0.132</td>
<td>0.084</td>
<td>0.106</td>
</tr>
<tr>
<td>Gentleness</td>
<td>0.169*</td>
<td>0.180*</td>
<td>0.061</td>
<td>0.321**</td>
<td>0.229</td>
<td>0.234</td>
<td>0.068</td>
<td>0.18</td>
<td>-0.055</td>
</tr>
<tr>
<td>Flexibility</td>
<td>0.194**</td>
<td>0.278****</td>
<td>0.014</td>
<td>0.26</td>
<td>0.368**</td>
<td>-0.21</td>
<td>0.166</td>
<td>0.249**</td>
<td>0.033</td>
</tr>
<tr>
<td>Patience</td>
<td>0.288***</td>
<td>0.268****</td>
<td>0.166*</td>
<td>0.349**</td>
<td>0.359**</td>
<td>0.08</td>
<td>0.239**</td>
<td>0.222*</td>
<td>0.19</td>
</tr>
</tbody>
</table>

***: \( P > 0.01 \); **: \( P < 0.05 \); *: \( P < 0.1 \). All correlations in the table are Spearman correlation coefficients. All significance levels indicated are two-tailed for the sake of comparability.

Table 7. Regressions of Four Personality Measures on Mean Contribution in Stage 1

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>t</th>
<th>p</th>
<th></th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>6.00</td>
<td>15.19</td>
<td>0</td>
<td>Constant</td>
<td>6.00</td>
<td>14.54</td>
<td>0</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.79</td>
<td>-1.17</td>
<td>0.25</td>
<td>Gender</td>
<td>-1.08</td>
<td>-1.51</td>
<td>0.13</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>1.23</td>
<td>1.58</td>
<td>0.12</td>
<td>Forgiveness</td>
<td>0.25</td>
<td>0.43</td>
<td>0.66</td>
</tr>
<tr>
<td>Gender x Agreeableness</td>
<td>3.89</td>
<td>2.68</td>
<td>0.01</td>
<td>Gender x Forgiveness</td>
<td>2.46</td>
<td>2.52</td>
<td>0.01</td>
</tr>
</tbody>
</table>

\( R^2 = 0.17 \)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>t</th>
<th>p</th>
<th></th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>5.98</td>
<td>14.74</td>
<td>0</td>
<td>Constant</td>
<td>5.98</td>
<td>14.93</td>
<td>0</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.58</td>
<td>-0.83</td>
<td>0.41</td>
<td>Gender</td>
<td>-0.55</td>
<td>-0.81</td>
<td>0.42</td>
</tr>
<tr>
<td>Flexibility</td>
<td>1.15</td>
<td>1.46</td>
<td>0.15</td>
<td>Flexibility</td>
<td>1.57</td>
<td>1.89</td>
<td>0.06</td>
</tr>
<tr>
<td>Gender x Flexibility</td>
<td>2.55</td>
<td>1.89</td>
<td>0.06</td>
<td>Gender x Flexibility</td>
<td>2.30</td>
<td>2.33</td>
<td>0.02</td>
</tr>
</tbody>
</table>

\( R^2 = 0.12 \)

Sample size: \( n = 108 \). All \( p \) values provided are two-tailed.
146. Jan Winiecki: Successes of Trade Reorientation and Trade Expansion in Post -
147. Jan Winiecki: Cost and Benefits of European Union’s Enlargement: a (largerly)
148. Alexander Kritikos: The Enforcement of Enviromental Policy under Incomplete
Information, Januar 2000.
149. Stefan Schipper und Wolfgang Schmid: Trading on the Volatility of Stock Prices,
151. Eberhard Stickel: Entrepreneur or Manager: Who really runs the Firm?, Februar
2000.
152. Wolfgang Schmid und Stefan Schipper: Monitoring Financial Time Series,
Februar 2000.
153. Wolfgang Schmid und Sven Knoth: Kontrollkarten für abhängige Zufallsvariablen,
Februar 2000.
154. Alexander Kritikos und Frank Wießner: Ein zweiter Kreditmarkt für eine zweite
Chance, Februar 2000.
155. Alexander Kritikos: A Discussion on the Viability of the Indenture Game, between
156. Claudia Kurz: Regional Risk Sharing and Redistribution by the Unemployment
Insurance: The Case of Germany, April 2000.
157. Friedel Bolle und Andreas Paul: Preventing International Price Discrimination – Are
158. Dorothea Baun: Operationalisierung der Determinanten von Impulskäufen –
159. Alexander Haupt: Environmental Policy and Innovations in Open Economies,
Juni 2000.
160. Jochen Hundsfelder: Lock-In-Effekte bei Gewinnen von Kapitalgesellschaften vor
und nach der geplanten Steuerreform, Juni 2000.
161. Alexander Kritikos und Friedel Bolle: Distributional Concerns: Equity or Efficiency
162. Sandra Große und Lars-Olaf Kolm: Anrechnung nach § 34c Abs. 1 oder Abzug
nach § 34c Abs. 2 EStG – Modellierung einer Entscheidungshilfe, August 2000.
163. Swantje Heischkel und Tomas Oeltze: Grundzüge des russischen
Umsatzsteuerrechts, August 2000.
165. Friedel Bolle und Alexander Kritikos: Reciprocity, Altruism, Solidarity: A dynamic
model, September 2000.
166. Jan Winiecki: An inquiry into the early drastic fall of output in post-communist

Eine Übersicht über die zwischen 1993 bis 1999 erschienenen Diskussionspapiere kann beim Dekanat der
Wirtschaftswissenschaftlichen Fakultät angefordert werden.


175. **Friedel Bolle**: If you want me, I don’t want you, December 2001.


178. **Andrea Gröppel-Klein und Dorothea Baun**: The more the better? – Arousing merchandising concepts and in-store buying behavior, Februar 2002.

179. **Yves Breitmoser**: Collusion and Competition in Laboratory Simultaneous Multiple-Round Auctions, Mai 2002.


182. **Friedel Bolle**: Altruism, Beckerian Altruism, or Intended Reciprocity? Remarks on an Experiment by Selten and Ockenfels, Mai 2002.

183. **Yves Breitmoser**: Subgame-Perfect Equilibria of Small Simultaneous Multiple-Round Auctions, Juni 2002.


189. **Friedel Bolle and Jessica Kaehler**: „Is there a Harmful Selection Bias when Experimenters Choose their Experiments?”, February 2003.


Friedel Bolle and Antje Baier: Cyclical Price Fluctuations caused by Information Inertia - Evidence from the German Call-by-Call Telephone Market, Januar 2005.


