Journal of Economic Behavior & Organization 81 (2012) 664–676

Contents lists available at SciVerse ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Temporal stability and psychological foundations of cooperation preferences

Stefan Volk[a,*], Christian Thöni[b], Winfried Ruigrok[c]

[a] University of Tübingen, Faculty of Business Administration and Economics, Melanchthonstrasse 30, 72074 Tübingen, Germany
[b] University of St. Gallen, Department of Economics, Varnbüelstrasse 14, 9000 St. Gallen, Switzerland
[c] University of St. Gallen, Research Institute for International Management, Dufourstrasse 40a, 9000 St. Gallen, Switzerland

ARTICLE INFO

Article history:
Received 1 February 2011
Received in revised form 7 October 2011
Accepted 13 October 2011
Available online 20 October 2011

JEL classification:
C91
C72
H41

Keywords:
Social preferences
Preference stability
Conditional cooperation
Free riding
Personality
Big-Five

ABSTRACT

A core element of economic theory is the assumption of stable preferences. We test this assumption in public goods games by repeatedly eliciting cooperation preferences in a fixed subject pool over a period of five months. We find that cooperation preferences are very stable at the aggregate level, and, to a smaller degree, at the individual level, allowing us to predict future behavior fairly accurately. Furthermore, our results provide evidence on the psychological foundations of cooperation preferences. The personality dimension 'Agreeableness' is closely related to both the type and the stability of cooperation preferences.

1. Introduction

In his influential essay on "The Methodology of Positive Economics" (1953, p. 4), Friedman explains that the task of economic theory "is to provide a system of generalizations that can be used to make correct predictions about the consequences of any change in circumstances".

In brief, an important aspect of economic theory is to make predictions, which are enormously facilitated if peoples' tastes are stable across time and circumstances. The assumption of stable preferences belongs accordingly to the core elements of economic theory, or, as Becker (1976, p. 5) puts it: "The combined assumptions of maximizing behavior, market equilibrium, and stable preferences, used relentlessly and unflinchingly, form the heart of the economic approach as I see it".

On the same page Becker also argues that "preferences are assumed not to change substantially over time". He explains that the assumption of preference stability both across different circumstances and over time "provides a stable foundation for generating predictions about responses to various changes, and prevents the analyst from succumbing to the temptation of simply postulating the required shift in preferences to "explain" all apparent contradictions to his predictions". Whether

* Corresponding author. Tel.: +49 7071 29 78179; fax: +49 7071 29 5534.
E-mail addresses: stefan.volk@uni-tuebingen.de (S. Volk), christian.thoeni@unisg.ch (C. Thöni), winfried.ruigrok@unisg.ch (W. Ruigrok).

0167-2681/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jebo.2011.10.006

the assumption of preference stability is reasonable for real world preferences is ultimately an empirical question. However, empirical evidence on preference stability is surprisingly scarce.

In this paper we present experimental evidence on the stability of social preferences, more precisely on the stability of preferences for cooperation in a social dilemma. We measure cooperation preferences by Fischbacher et al.'s (2001, henceforth FGF) variant of the four player one-shot public goods game and repeat the measurement with the identical subject pool after two and a half and five months. The FGF method has the advantage of measuring cooperativeness to a large degree independent of subjects' beliefs, which are presumably more volatile than preferences.[1] Furthermore, we also present evidence on the psychological foundations of cooperation preferences, more specifically on the effects of individual differences in personality on these preferences. Our experimental setup enables us to investigate (i) aggregate-level and individual-level stability of cooperation preferences over time and (ii) the relationship between stable personality traits and cooperation preferences.

Our paper contributes to a nascent body of literature on the stability of social preferences. However, the existing literature has largely focused on the distribution of heterogeneous preference types across different experiments at the aggregate level.[2] The few studies that have examined individual-level stability can be divided into two categories.

The first category includes studies investigating the consistency and stability of social preferences across different games or variants of the same game in a given experimental session. Andreoni and Miller (2002) and Fisman et al. (2007) apply the axioms of revealed preferences to decisions observed in modified dictator games. They find that subjects' choices can be rationalized by a well-behaved utility function. Fischbacher and Gächter (2010) elicit individual cooperation preferences in a strategy method experiment and observe contributions in ten consecutive one-shot games in the direct response mode. They find that data on individual cooperation preferences allow to predict the development of contributions over time very accurately.[3] Blanco et al. (2011) investigate the stability of preferences across different games by observing the same sample of subjects in a number of different one-shot games. They find that about one third of their subjects exhibit stable preferences in the sense that they consistently follow a plausible behavioral norm.[4]

The second category of research includes a very small number of studies examining longitudinal stability of social preferences over time. Muller et al. (2008) elicit subjects' cooperation preferences repeatedly in a series of five consecutively played two-stage public goods experiments within one session. They find that 37 percent of their 60 subjects remain in the same preference category for all five measurements. Studies allowing for more time between the measurements are particularly scarce.[5] Brosig et al. (2007) implement sequential prisoner's dilemma games and repeat the initial experiment with the same subject pool two times with one month in between each repetition. They find that 43 percent of the subjects choose the same response in all three waves, and all of these subjects act consistently selfish. On average they observe a decay of cooperative behavior over time.

In contrast to Brosig et al. (2007), we observe highly stable aggregate results. The distribution of cooperation preferences is basically unchanged across time. At the individual level we observe a more diverse picture. If we classify our subjects into three categories (Conditional Cooperators, Free Riders and Others) we find that about two thirds of the subjects remain in the same category for two consecutive measurements. Half of the subjects remain in the same category for all three measurements.

While the temporal stability of cooperation preferences is the main focus of the present paper, we are also exploring the psychological foundations of preference heterogeneity and stability. The idea to include psychological variables into the analysis of cooperation preference stability came from a related paper (Volk et al., 2011), which is based on the same data as this paper, and which served to some extent as a preparatory analysis for the present paper. In Volk et al. (2011) the main focus is on how personal values and personality traits are related to one another and how they jointly impact cooperation preferences. This analysis is important because one of the problems inherent in the study of psychological correlates of economic behavior is the large number of conceivably important psychological constructs to pick from. Two particularly interesting constructs are personality traits and personal values as they represent the difference between innate characteristics representing the nature of an individual (i.e., personality traits) and socially learned characteristics resulting from the interaction of nature and nurture (i.e., personal values) (see for example Olver and Mooradian, 2003; Parks and

[1] There is of course no way to measure preferences directly. In a strict sense it is thus impossible to test preference stability. We also cannot identify the proximate mechanism which causes subjects to contribute, be it interdependent preferences or reciprocity (see Cox, 2004; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006; Fehr and Schmidt, 1999). The FGF design offers a way to measure contribution strategies in public goods games. A large variety of social preferences give rise to identical contribution strategies. Thus, in a strict sense, all we can show is the stability of contribution strategies over time. For many applications (e.g. predicting future behavior in similar situations) it is sufficient to know that contribution strategies are stable.

[2] For experiments using the FGF design see Kocher et al. (2008), Herrmann and Thöni (2009), and Thöni et al. (2009). Homo/heterogeneity between subject pools in repeated public goods games are studied in Gächter et al. (2010).

[3] Burlando and Guala (2005) and Gächter and Thöni (2005) investigate preference stability in a setting where they regroup subjects according to their cooperation preferences. They use initial public goods games to identify different types and form groups of alike subjects and observe contributions in repeated public goods games.

[4] See also De Oliveira et al. (2009). A related strand of literature investigates whether different elicitation methods (strategy method vs. direct response method) lead to identical inference about subjects' preferences, see Brandts and Charness (2011) for an overview or Fischbacher and Gächter (2009) for the FGF design.

[5] Horowitz (1992), Andersen et al. (2008) and Zeisberger et al. (in press) investigate the stability of risk preferences over time. Meier and Sprenger (2010) present data on the stability of time preferences.

Guay, 2009). The results of Volk et al. (2011) suggest that while personality traits and personal values are both indicative of individual preferences for cooperation, personality traits are the stronger predictor. In the present paper we therefore focus on personality traits. Our results indicate that the personality dimension Agreeableness, which is related to one's tendency to be cooperative rather than competitive (Costa and McCrae, 1992), appears to be part of the psychological foundations of cooperation preference heterogeneity and stability. We find that Free Riders and Conditional Cooperators are characterized by different levels of Agreeableness and that higher levels of Agreeableness are associated with more stable preferences for conditional cooperation, while lower levels are associated with more stable preferences for free riding.

The latter findings contribute to an emerging research stream that seeks to identify "personality and attitude variables that allow the classification of subjects into different "types" whose decision behavior in social dilemmas may be described by alternative models" (Rapoport and Suleiman, 1993, p. 193).[6] Indeed, economists and psychologists are increasingly recognizing the potential of incorporating insights from personality psychology into economic models of preference formation and stability. Ferguson et al. (2011) argue that the overlap between personality traits and preferences is a key research issue in both economics and psychology. They also point out that "[e]conomic models of preferences will benefit from studies on cross-situational consistency and temporal stability" (p. 202). In a comprehensive review of the relationship between personality traits and economic preferences, Almlund et al. (forthcoming, p. 157) conclude that personality psychology "promises to provide a deeper understanding of conventional economic preference parameters and how they arise. Unfortunately, at the time of this writing, this promise remains unfulfilled". In this paper we attempt to contribute to this line of research.

The remainder of the paper is organized as follows. In Section 2 we describe our experimental design and our questionnaire, in Section 3 we present the results and Section 4 concludes.

2. Methods

2.1. Experimental design and procedures

The data reported in this paper come from an experimental part and a questionnaire part. In the experimental part we used the design introduced by FGF to measure cooperation preferences in public goods games. In order to analyze whether cooperation preferences are stable over time, we replicated the protocol used by FGF and repeated our initial study (Wave 1) in a random matching mode two and a half (Wave 2) and five months (Wave 3) after the first study by inviting the same participants back to the laboratory. All sessions used an identical protocol in which subjects were randomly assigned to groups of four members. Group compositions were unknown to the subjects and not revealed after the experiments. The basic decision situation was a standard linear public goods game. Each subject received an endowment of 20 tokens and chose a contribution $c_i \in \{0, 1, \ldots, 20\}$ to a linear public good with a marginal per capita return of 0.4. The payoff function was given by

$$\pi_i = 20 - c_i + 0.4\sum_{j=1}^{4} c_j,$$

where the public good is equal to the sum of the contributions of all four group members.

In the experiment subjects had to make two types of contribution decisions, an 'unconditional' and a 'conditional' contribution to the public good. The unconditional contribution was a single decision about how many of the 20 tokens to invest into the public good. For the conditional contribution, subjects had to fill in a table showing the 21 possible average contribution levels of the other three group members (rounded to integers). They were asked to state for each of the 21 possibilities their corresponding contribution. After all subjects had made both types of decisions, a random mechanism determined which of the two decisions became outcome relevant. The random mechanism (throw of a die) selected in each group one subject that contributed according to his or her 'contribution table', while the other three group members contributed according to their unconditional contribution.

We conducted our experiments in the computer lab of a European university, using the software z-Tree (Fischbacher, 2007). We observed the decisions of 72 subjects at times 1 and 2. Four of these did not return for the third study leaving us with 68 subjects with complete information at all three points in time. Subjects were informed in Wave 1 that they would participate in three identical experiments over the course of five months. They were also informed that group compositions would change randomly in each wave. In the experiments participants were randomly seated at separated computer

[6] The relation between personality measures and other-regarding, prosocial behavior has been studied in public goods games (Fleming and Zizzo, 2011; Kurzban and Houser, 2001; Perugini et al., 2010; Skatova and Ferguson, 2011; Volk et al., 2011), dictator games (Ben-Ner et al., 2004a,b; Ben-Ner and Kramer, 2011), ultimatum games (Brandstätter and Königstein, 2001) and a series of distribution games (Bartling et al., 2009). Ashton et al. (1998) and Hirsh and Peterson (2009) study the correlation between personality measures and nonincentiviced measures for cooperation. Boone et al. (1999, 2002) correlate personality measures like 'Locus of control' with cooperation in prisoners' dilemma games. Gunnthorsdottir et al. (2002) and Meyer (1992) investigate the relationship between Machiavellianism and behavior in the trust game and the ultimatum game, respectively. For the relation between survey measured trust and experimentally measured cooperative behavior see also Gächter et al. (2004).

terminals, received written instructions and had to answer a number of control questions. At the end of the experiments the participants were informed about their final payoff and paid privately. On average, subjects earned about US$ 13 in each of the three experiments.[7]

2.2. Questionnaire

We assessed individual differences in personality traits by a measure of the Big-Five model of personality which was administered after the experiment of Wave 1. Personality traits have an estimated annual stability coefficient of 0.98, indicating high stability over lifetime (Conley, 1985).[8] Research has accordingly shown that subjects completing a personality questionnaire more than once will tend to obtain highly similar scores (e.g., McCrae and Costa, 1990). In order to reduce the potential for transient measurement error (e.g., Chmielewski and Watson, 2009) resulting from participant frustration and carelessness associated with completing the same personality questionnaire repeatedly, we measured personality characteristics in Wave 1 and did not repeat the measurement in the subsequent waves.

The Big-Five model is the standard trait framework for research in personality (Goldberg, 1993; John and Srivastava, 1999) and has enjoyed increasing popularity across a wide variety of disciplines including economics.[9] The model specifies that five overarching dimensions account for the biggest part of between-subject variation in stable personality traits. These dimensions are Extraversion (sociable, active, and energetic), Agreeableness (cooperative, considerate, and trusting), Conscientiousness (dependable, organized, and persistent), Emotional Stability (calm, secure, and unemotional) and Openness to Experience (imaginative, intellectual, and artistically sensitive). While we measured all five dimensions, one of them, Agreeableness, is particularly relevant for our study of cooperation preferences. Agreeableness is centrally related to one's tendencies to strive for cooperation rather than competition (Costa and McCrae, 1992). Moreover, individuals high on Agreeableness are generally more inclined to forgo self-interests in favor of collective interests than individuals low on Agreeableness (e.g., Buss, 1991; Koole et al., 2001). Graziano and Eisenberg (1997) therefore conclude that Agreeableness is the personality dimension most closely related to prosocial behavior. We therefore expected Agreeableness to be positively correlated with subjects' cooperativeness in our experiments.

We elicited personality traits by a Ten-Item Personality Inventory (TIPI), which includes a subscale for each of the five personality dimensions (Gosling et al., 2003). Each subscale is composed of two items and each item contains a pair of two trait descriptors. For example, the first item is: "*I see myself as* extraverted, enthusiastic". Subjects have to rate on a 7-point scale ranging from 1: 'disagree strongly' up to 7: 'agree strongly' the extent to which the pair of traits applies to them. The ten items are 'extraverted, enthusiastic' and 'reserved, quiet' (reverse-scored) for the Extraversion subscale; 'sympathetic, warm' and 'critical, quarrelsome' (reverse-scored) for the Agreeableness subscale; 'dependable, self-disciplined' and 'disorganized, careless' (reverse-scored) for the Conscientiousness subscale; 'calm, emotionally stable' and 'anxious, easily upset' (reverse-scored) for the Emotional Stability subscale; and 'open to new experiences, complex' and 'conventional, uncreative' (reverse-scored) for the Openness to Experience subscale.[10]

Gosling et al. (2003) reported acceptable psychometric properties for the TIPI with test-retest reliabilities over a period of six weeks of 0.77 for Extraversion, 0.71 for Agreeableness, 0.76 for Conscientiousness, 0.70 for Emotional Stability and 0.62 for Openness with a mean of 0.72 over all five subscales, indicating that the scale provides a stable measure of personality over time.

3. Results

We organize our results as follows: We start by conducting an instrument check of our experiment and look at aggregate stability and individual stability. We then conduct an instrument check of our personality measure before we explore the impact of the Big-Five personality dimensions on preference heterogeneity and preference stability. Finally, we investigate the predictive power of preference stability. Specifically, we use the data from the first and second wave to predict individual behavior in the third wave and evaluate the prediction with the true data.

3.1. Instrument check public goods game

Since we use the FGF design, we start by checking whether we replicate previous findings from the literature in our Wave 1 experiment. Following the classification of preference types proposed by FGF, we divide our subjects into four categories.

---

[7] In all three waves subjects played a one-shot public goods game and a one-shot public goods game with punishment prior to the experiment reported here. Subjects did not, however, receive any information about other subjects' decisions nor about their payoff. At the end of the experiments they were informed only of their aggregate payoff.

[8] This annual stability coefficient is a corrected estimate based on a wide array of inventory and rating studies.

[9] See, e.g., Ameriks et al. (2007), Bartling et al. (2009), and Lo et al. (2005).

[10] We administered the original well-validated English version of the TIPI (Gosling et al., 2003, p. 525). While our participants were non-native English speakers, their English skills were sufficient to properly respond to the English TIPI. The mean scores, standard deviations, inter-item correlations and Cronbach's alphas for the five TIPI subscales obtained in the present study are very similar to those reported in three previous studies using the TIPI with native English speakers (Donnellan et al., 2006; Ehrhart et al., 2009; Gosling et al., 2003), indicating that the English TIPI items performed well in our setting.

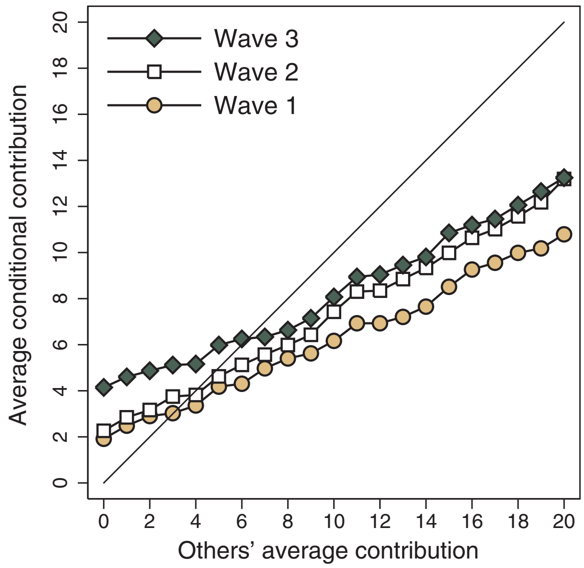

Fig. 1. Average conditional cooperation for the three waves.

All subjects with a contribution schedule that has a significant positive slope ($p < 0.01$, Spearman rank correlation) or shows a monotonically increasing pattern are classified as 'Conditional Cooperators'. 'Free Riders' contribute nothing in any case. Subjects who increase their contribution in the contribution schedule up to some maximum and decrease it thereafter are called 'Hump-shaped'. The remaining subjects fall into the category 'Others'. We (FGF) observe 58.3 (50) percent Conditional Cooperators, 25.0 (29.5) percent Free Riders, 2.8 (13.6) percent Hump-shaped and 13.9 (6.8) percent Others.[11] We run a χ^2 test on the joined data to check whether our distribution of types systematically differs from the one observed by FGF, and indeed the differences are borderline significant with $\chi^2(3) = 6.44$; $p = 0.092$. A closer look at the percentage numbers reveals that the greatest difference between our data and the FGF data is the smaller number of Hump-shaped patterns we observe. This is not completely unexpected, as other replication studies of the FGF design (e.g., Herrmann and Thöni, 2009; Kocher et al., 2008; Thöni et al., 2009) also find a lower percentage of Hump-shaped patterns. Due to the infrequent occurrence in our data we include the Hump-shaped patterns into the category Others. If we compare our data to FGF using this reduced classification we observe no significant differences ($\chi^2(3) = 0.771$; $p = 0.680$). We conclude that, with regard to the most interesting categories, our Wave 1 experiment produced very similar results in comparison to the previous literature.

Remember that our subjects also made unconditional contribution decisions. Our subjects made average unconditional contributions of 9.1 tokens (45.7 percent, $N = 72$) in Wave 1, 10.1 tokens (50.6 percent, $N = 72$) in Wave 2 and 10.3 tokens (51.3 percent, $N = 68$) in Wave 3. The differences between the three waves are insignificant according to Wilcoxon signed rank tests. The corresponding p-values are $p = 0.494$ ($p = 0.575$, $p = 0.987$) for the comparison between Wave 1 and Wave 2 (Wave 1 and Wave 3, Wave 2 and Wave 3). In the FGF experiment subjects made a slightly lower average unconditional contribution of 6.7 tokens (33.5 percent of the endowment, $N = 44$).[12]

3.2. Aggregate stability

Before we classify our data from the three waves into categories we look at the overall conditional cooperation scheme. Fig. 1 shows the average conditional cooperation schemes for the three waves. The horizontal axis depicts the average contribution of the other three subjects, on which a subject can condition the own contribution. The vertical axis shows the average of all conditional contribution decisions in a wave.

The curves for the three waves have a similar positive slope, indicating that on average subjects are Conditional Cooperators. There seems to be a slight positive trend, i.e., the level of the average conditional contribution is higher in later waves. We therefore test for systematic differences between the three times of measurement using the average conditional cooperation of a subject in a wave as the test variable. Our results show no significant differences between the three waves.[13]

[11] $N = 72$ for our Wave 1 experiment and 44 for the FGF study. The corresponding frequencies for our study (FGF) are: 42 (22) Conditional Cooperators, 18 (13) Free Riders, 2 (6) Hump-shaped and 10 (3) Others.

[12] The means of the unconditional contributions of the different types of subjects in Wave 1 (Wave 2; Wave 3) are as follows: Conditional Cooperators: 10.62 (11.98; 11.18), Free Riders: 3.72 (1.57; 0) and Others: 12.08 (12.73; 16).

[13] We use Wilcoxon signed rank tests to test for differences in the average conditional cooperation across waves. The corresponding p-values are $p = 0.182$ ($p = 0.113$; $p = 0.767$) for the comparison between Wave 1 and Wave 2 (Wave 1 and Wave 3; Wave 2 and Wave 3).

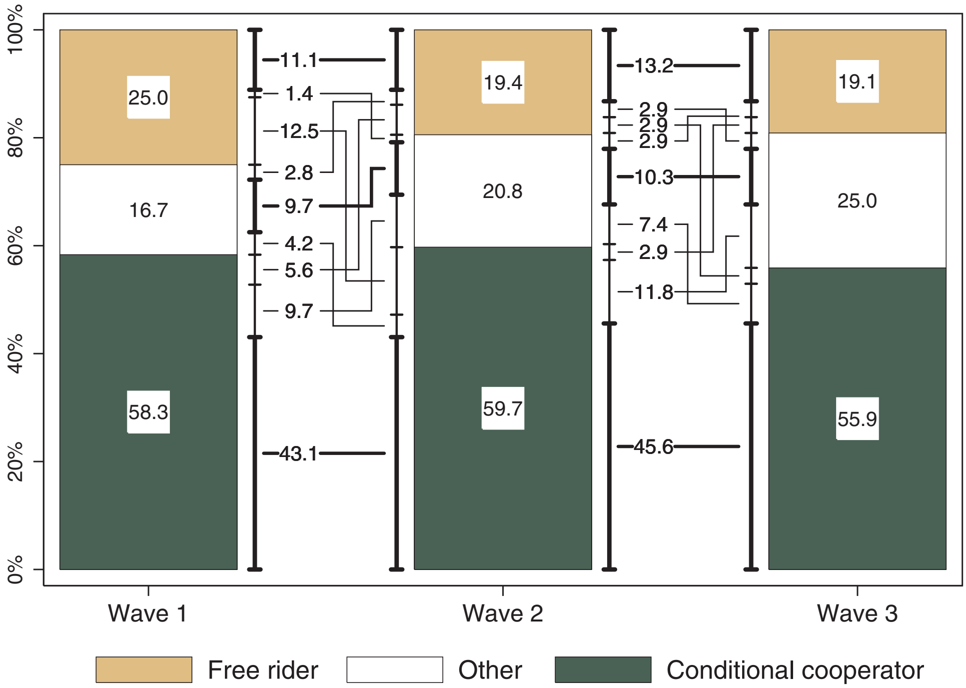

Fig. 2. Distribution of types in the three waves and transitions between the times of measurement. All numbers denote percentages of the whole population in a given wave.

Overall, conditional contributions are thus relatively similar over the course of five months in the same sample of subjects. Let us now check whether the same holds true for the distribution of types in the population.[14] The three bars in Fig. 2 show the frequency of Conditional Cooperators, Free Riders and Others in the three waves of our experiment (ignore for the moment the dizzying lines between the bars). The distribution of types is apparently very similar across the three times of measurement. Free Riders account for between 19 percent (Wave 3) and 25 percent (Wave 1) of the population, while Conditional Cooperators account for between 56 percent (Wave 3) and 60 percent (Wave 2). The distribution of preference types is not significantly different across the three times of measurement ($\chi^2(4) = 1.994$; $p = 0.737$). This finding in combination with the observation that the average conditional cooperation schemes are not significantly different across the three waves (Fig. 1) is in contrast to the results of Brosig et al. (2007) who report a decay of cooperation in the later measurements and speculate that their subjects learned the Free Rider strategy over time. While the learning hypothesis can neither explain our observations nor the observations of Muller et al. (2008), our data point to a different explanation for preference stability/instability at the individual level. As we will report in the next sections, our findings suggest that there is heterogeneity of individual behavior in terms of both preference types and preference stability (i.e., some individuals seem to have stable preferences, while others do not), which can be linked to individual differences in dispositional personality traits.

To conclude, our observations show high aggregate stability over time. Observed changes in the distribution of types as well as in average conditional contribution schemes are both small and insignificant. In all three waves we observe shares of Conditional Cooperators (56–60 percent) and Free Riders (19–25 percent), which are comparable to the shares reported in a number of previous studies. These studies have found that between 50 and 60 percent of people can be classified as cooperators, while between 20 and 30 percent are Free Riders (e.g., Fischbacher and Gächter, 2010; Kurzban and Houser, 2005). In a next step, we investigate whether this aggregate stability is caused by stability of the contribution patterns at the individual level.

3.3. Individual stability

In order to test for individual stability, we first check whether subjects classified into one of the three classes in the first (second) wave remain in that class in the second (third) wave. With 72 subjects in Waves 1 and 2 and 68 subjects in Wave 3, there are 140 opportunities for individual preferences to change type between two times of measurement. In 93 cases (66.4

[14] Note that the classification into types is only partially sensitive to changes in preferences, because the class Conditional Cooperators allows for many different degrees of cooperativeness. However, we think that for most applications (like e.g. calibrating theories of social preferences) it is sufficient to separate distinct classes of behavioral patterns and abstract from the heterogeneity within class.

Table 1
Relative frequencies of all 27 possible combinations of types in the three waves. Bold numbers show observations on the main diagonal, i.e., the percentage of observations that are of the same type in all three waves. Numbers in parentheses denote expected frequencies of combinations of types in all three waves if transitions are purely random; all numbers show percentages of the total of 68 observations.

Wave 3		Conditional Cooperator (CC)			Free Rider (FR)			Other (OT)		
Wave 2		CC	FR	OT	CC	FR	OT	CC	FR	OT
Wave 1	CC	**35.3** **(20.3)**	1.5 (6.4)	2.9 (6.9)	0.0 (6.9)	1.5 (2.2)	1.5 (2.4)	8.8 (9.1)	2.9 (2.9)	5.9 (3.1)
	FR	7.4 (8.4)	1.5 (2.7)	1.5 (2.9)	2.9 (2.9)	**10.3** **(0.9)**	0.0 (1)	1.5 (3.8)	0.0 (1.2)	0.0 (1.3)
	OT	2.9 (5)	0.0 (1.6)	2.9 (1.7)	0.0 (1.7)	1.5 (0.5)	1.5 (0.6)	1.5 (2.2)	0.0 (0.7)	**4.4** **(0.8)**

percent) subjects do not change their type between two points in time. Preference stability is somewhat stronger between Wave 2 and Wave 3 (69.1 percent vs. 63.9 percent between Waves 1 and 2).[15] Given that we have three measures for each subject we can also check whether a subject belongs to the same type in all three waves. Table 1 shows the frequency of all 27 combinations of types in the three waves. Bold numbers indicate the percentage of observations on the main diagonal which correspond to stable types in all three waves. Numbers in parentheses indicate the expected frequency which would result if types were drawn randomly in each wave from a distribution of types as observed in the respective wave.

Exactly half of the 68 subjects we observe three times are classified as the same type in all three waves, 35.3 percent display stable Conditional Cooperation preferences, 10.3 percent stable Free Rider preferences, and 4.4 percent are consistently classified as Others. Table 1 shows that for each type observed in Wave 1 it is clearly more frequent that the subject remains of the same type in later waves than any other combination of types. Comparing the actual frequencies to expected frequencies shows that stable types are much more likely than expected whereas most other combinations occur less frequent in the data than expected under the assumption of random types.

This indicates that we can clearly refute the hypothesis that types are random. To provide a test for this we simulate the stability of types under the assumption that each player randomly picks a type in Wave 2 and Wave 3 with a probability that equals the observed relative frequency of the three types in our data. In the simulation we observe that in 21.3 percent of the cases subjects are of the same type in two adjacent waves (100 runs, standard deviation 3.5 percent points). The hypothesis of random types can be rejected at any reasonable level of significance.[16]

What are the patterns of the type changes? The three bars in Fig. 2 show the fractions of Conditional Cooperators, Free Riders and Others at the three points in time. The numbers between the bars show the fraction of subjects that remains in a class or changes to another class in the subsequent wave. All numbers indicate the percentage of subjects relative to the whole population. For example, 58.3 percent of the subjects are classified as conditionally cooperative in Wave 1. About three fourths of these (and 43.1 percent of all subjects) remain Conditional Cooperators in Wave 2, while 5.6 percent of all subjects change from Conditional Cooperator to Free Rider from Wave 1 to Wave 2 and 9.7 percent of all subjects move from Conditional Cooperator to the category Others. Bold lines correspond to observations that remain in a category.

Overall, Fig. 2 shows that the type of the Conditional Cooperator is clearly the most stable category, with three fourth of the subjects remaining in that class between two waves. Surprisingly, the category of Free Riders seems to be rather unstable, especially between Wave 1 and Wave 2. Among the 25 percent Free Riders in Wave 1, more than half change their type in Wave 2, most of those who change join the conditionally cooperative group. The fraction of Free Riders which stay in that category increases to almost 70 percent when we compare Wave 2 and Wave 3. However, the group of conditionally cooperative subjects is still more stable with about 77 percent of the subjects remaining in the same class between Wave 2 and Wave 3.

In conclusion, our results indicate that, over the course of five months, there is considerable stability of the distribution of types both at the individual and at the aggregate level. In a next step, we will use our questionnaire data to explore the impact of the Big-Five personality dimensions on preference heterogeneity and preference stability.

3.4. Personality and cooperation preferences

We first conduct an instrument check of our personality measure by running a principal components analysis with varimax rotation on the 10 items of the TIPI scale.[17] The analysis suggests a five factor solution which explains 74 percent of the variance. Factor loadings are all above 0.60 and all items load on their hypothesized factors.

---

[15] This result is remarkably similar to the findings of Muller et al. (2008), who examine 60 subjects playing a comparable game in 5 repetitions within the same experimental session. In their design there are 240 opportunities for individual preferences to change and in 66.7 percent of cases subjects remain in the same category in the next game.

[16] Another measure of the stability of types between waves is Cohen's κ (or Fleiss's κ for all three waves): In our data we find $\kappa_{12} = 0.365$, $\kappa_{13} = 0.288$, $\kappa_{23} = 0.463$, $\kappa_{123} = 0.371$ (subscripts denote waves).

[17] Principal components analysis is a factor analysis technique that is commonly used as an instrument check in psychometric research to verify the factor structure (in case of the TIPI the hypothesized five factor structure) of psychological measures. Varimax rotation is part of the factor analysis procedure and is done to make the factor solution more interpretable. For more details see e.g. Nunnally and Bernstein (1994).

Table 2
Multinomial logit estimate for type of conditional cooperation scheme in Wave 1. Omitted case is Conditional Cooperator. Independent variables are the five personality dimensions, demeaned and standardized. Standard errors in parentheses.

	Multinomial Logit	
	Free Rider	Other
Extraversion	−0.363	0.569
	(0.363)	(0.429)
Agreeableness	−0.898**	−0.924**
	(0.390)	(0.420)
Conscientiousness	−0.020	0.384
	(0.329)	(0.415)
Emotional Stability	0.076	−0.241
	(0.342)	(0.405)
Openness to Experience	0.782**	−0.235
	(0.395)	(0.360)
Constant	−1.060***	−1.468***
	(0.340)	(0.402)
Model χ^2	20.1	
Prob > χ^2	0.028	
Pseudo *R*-squared	0.146	
N	72	

** Significance at 5 percent.
*** Significance at 1 percent.

In this section, we investigate whether the Big-Five personality dimensions are related to cooperation preferences. The analysis reported here complements a related analysis in Volk et al. (2011). We use a multinomial logit model with the Big-Five personality dimensions as explanatory variables to explain the observed cooperation preference type in Wave 1. Table 2 shows the results of these estimates. The omitted case is the Conditional Cooperator type. Table 2 shows that two Big-Five dimensions are significantly related to the observed type in Wave 1, Agreeableness and Openness to Experience.

Subjects scoring high on Agreeableness are significantly less likely to be classified as Free Riders compared to Conditional Cooperators. The same is true for the category Others. Furthermore, scoring high on Openness to Experience is associated with a higher probability of being classified as a Free Rider compared to being classified as a Conditional Cooperator.

What is the size of the personality effect? We use multinomial logit models to predict probabilities of the three types dependent on the Agreeableness score in all three waves. Panels A to C in Fig. 3 present the results. The horizontal axis shows the range of Agreeableness scores we observe in our experimental data (for the distribution of scores see Fig. A2 in Appendix A). To facilitate the interpretation we de-mean and standardize the values of Agreeableness. The size of the estimated effect is quite remarkable. In Wave 1, a subject scoring 2.7 standard deviations lower than the average on Agreeableness has an estimated probability of about 13 percent to be classified as Conditional Cooperator. For the average subject the probability is about 63 percent and increases up to 86 percent for a subject scoring 1.4 standard deviations higher than the average on Agreeableness. Conversely, the estimated probability of being classified as Free Rider drops from 51 percent to a mere 8 percent if we compare the least to the most agreeable subject in our sample.

What about time effects? We can use our personality measures elicited in the experiment of Wave 1 to predict the type in Wave 2, or Wave 3. Panels B and C in Fig. 3 show the results. It turns out that the influence of Agreeableness on the types Free Rider and Conditional Cooperator is strong and significant in all three waves. The negative connection between Agreeableness and the type Other observed in Table 2 seems to be reversed in the subsequent waves. However, the effect is not significant. The positive influence of Openness to Experience on the category Free Rider is insignificant in Wave 2 and only marginally significant in Wave 3.

How are personality measures associated with preference stability over time? Given that we have three types of cooperation preferences and three times of measurement, there are ten possible combinations of types. A subject who is classified as Conditional Cooperator in all three waves is categorized as *ConsistentCC*. Likewise, we observe three times Free Riders (*ConsistentFR*). Subjects who are classified as either Conditional Cooperator or Free Rider in two out of the three waves are categorized as *2timesCC* and *2timesFR*, respectively. All remaining combinations are in the same group (*All other*).[18] We run a multinomial logit model for these five categories. Like before, only the personality dimension Agreeableness shows significant results. Panel D of Fig. 3 shows the predicted probabilities dependent on Agreeableness. Clearly, the probability of *ConsistentFR* is now very strongly connected to the Agreeableness measure. It ranges from 87 percent for the lowest observed Agreeableness to virtually zero for an Agreeableness score that is a standard deviation above the mean. The category *2timesFR* shows a similar pattern. On the other end of the scale we observe that the probability of *ConsistentCC* rises from virtually zero to 43 percent for the highest Agreeableness score observed in our sample. The category *2timesCC* displays a similar pattern. Finally, unlike in Panel A, the remaining observations are more likely to be associated with higher scores of Agreeableness.

[18] We do not introduce a separate category for subjects classified as Other in all three waves because (i) only three subjects do so and (ii) it is unclear how we should interpret a pattern that is consistently inconsistent.

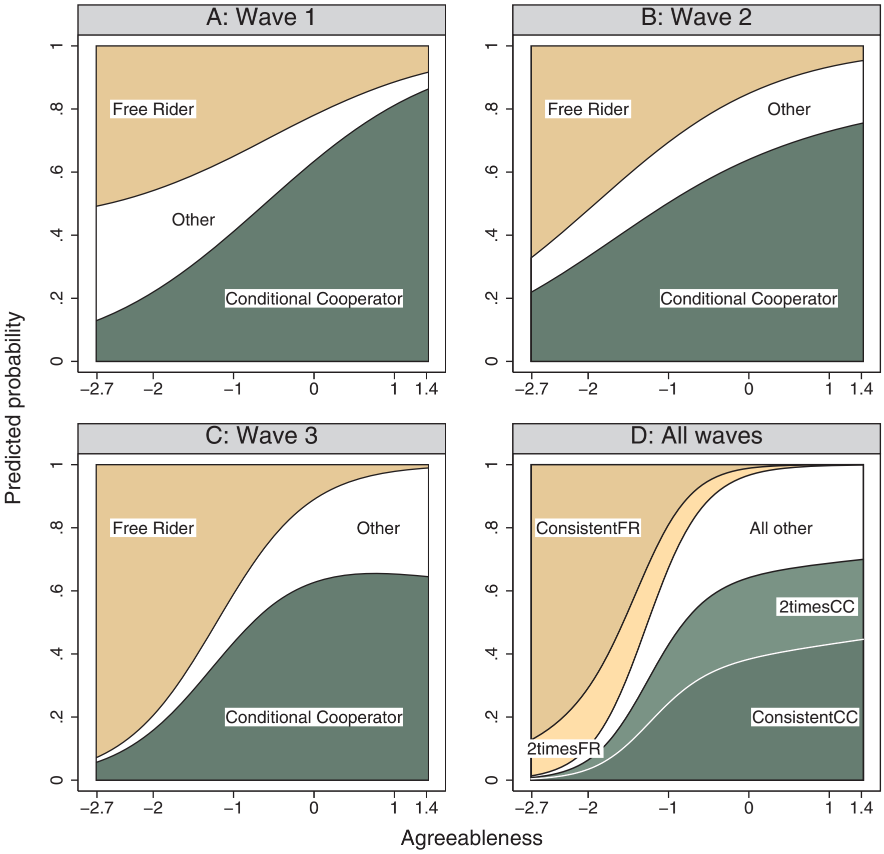

Fig. 3. Panels A–C: estimated probability of being classified as either 'Conditional Cooperator', 'Free Rider', or 'Other' dependent on the Big-Five dimension Agreeableness (demeaned and normalized) in Waves 1, 2 and 3. Panel D: estimated probability of being classified into one of the five classes during all three waves of the experiment dependent on Agreeableness. 'Consistent FR (CC)' are subjects classified as Free Rider (Conditional Cooperator) in all three waves; 2timesFR (CC) are subjects classified as Free Rider (Conditional Cooperator) in two out of the three waves; 'All other' denotes all remaining patterns.

Taken together, our results suggest that the personality dimension Agreeableness can contribute to our understanding of the psychological foundations of cooperation preference heterogeneity and stability. Concerning preference heterogeneity, we find that Free Riders and Conditional Cooperators are characterized by different levels of Agreeableness. Regarding preference stability, we find that higher (lower) levels of Agreeableness are associated with more stable preferences for conditional cooperation (free riding). These findings are consistent with the limited available evidence on the relationship between cooperativeness in economic experiments and the Big-Five dimension Agreeableness. For example, Agreeableness has been shown to be related to higher contributions in public goods games (Perugini et al., 2010), more cooperative choices in prisoner's dilemma games (Pothos et al., 2011) and higher donations in dictator games (Ben-Ner et al., 2004a). Furthermore, Bartling et al. (2009) measure competitiveness in a sample of mothers of preschool children by confronting them with the choice between competing in a tournament or receiving a piece rate for a real effort task. They find that subjects scoring high on the personality dimension Agreeableness are less likely to self-select into competition, i.e., into the tournament.

3.5. Predicting types

In the beginning of the paper we have argued that an important aspect of economic theory is to make predictions and that making predictions is enormously facilitated if peoples' tastes are stable. So far we have shown that cooperation preferences are stable at the aggregate level. If we are interested in aggregate outcomes, a prediction that simply extrapolates the observed share would be fairly accurate. At the individual level we find that about one third of the subjects changes preference type between two waves. What does this mean for the predictive power of potential models predicting individual types? In the following, we use the data from Wave 1 and Wave 2 to predict a subject's type in Wave 3. We then evaluate the predictions with the data from Wave 3. We calculate the accuracy for different methods of predicting individual types in Wave 3, making increasingly use of the information we have about our subjects from Waves 1 and 2. The accuracy of a model is simply the number of subjects for whom we correctly predict the type in relation to all subjects.

A first benchmark is a prediction that completely ignores preference stability and does not take information from Waves 1 and 2 into account. Given that there are three types of cooperation preferences we would make the uneducated guess of

Table 3
Confusion matrices for four models to predict types in Wave 3 based on observations in Wave 1 and Wave 2.

	Predicted type		Observed type			Performance measures		
			CC	FR	OT	Accuracy (%)	Errors (%)	
			38	13	17		Type I	Type II
Model 1 Covariates: CC_{t-1}, FR_{t-1}	CC	54	**33**	11	10		13.2	70.0
	FR	0	0	**0**	0	58.8	100.0	0.0
	OT	14	5	2	**7**		58.8	13.7
Model 2 Covariates: Big-Five	CC	60	**38**	6	16		0.0	73.3
	FR	8	0	**7**	1	66.2	46.2	1.8
	OT	0	0	0	**0**		100.0	0.0
Model 3 Covariates: CC_{t-1}, FR_{t-1}, Big-Five	CC	45	**32**	4	9		15.8	43.3
	FR	9	1	**7**	1	67.6	46.2	3.6
	OT	14	5	2	**7**		58.8	13.7
Model 4 Covariates: CC_{t-1}, FR_{t-1}, Agreeableness, interaction	CC	46	**33**	3	10		13.2	43.3
	FR	8	0	**8**	0	70.6	38.5	0.0
	OT	14	5	2	**7**		58.8	13.7

predicting each of the three types with equal probability. In this case we would predict the correct type in 33.3 percent of the cases (in expectation).[19]

If we only consider information about the aggregate outcome in Waves 1 and 2 to predict individual behavior in Wave 3, our best estimate for a subject's type would be the type that occurs most frequently, which is the Conditional Cooperator. In this case we would be correct for all Conditional Cooperators and wrong for all other types. Thus our prediction would be correct in 55.9 percent of the cases, which is the fraction of Conditional Cooperators observed in Wave 3.

In a next step we make use of individual information from Waves 1 and 2. A naïve prediction would assign each subject the type observed before. If we use the type observed in Wave 2 we reach a correct prediction in 69.1 percent of the cases (see Fig. 2). However, one could as well take the type observed in Wave 1 as prediction for the Wave 3 data, in which case the prediction would be correct in 58.8 percent of the cases. Ex ante (i.e., not knowing the outcome in Wave 3) we would probably not have been satisfied with both approaches, because they do not make use of all information available.

A natural way of making use of the data from both waves would then be to estimate a model explaining types in Wave 2 by information from Wave 1 and then use this model to predict types in Wave 3. Firstly, we run a multinomial logit model explaining types in Wave 2 with two dummies for a subject's type in Wave 1 as explanatory variables (CC_{t-1} and FR_{t-1}). This model allows predicting probabilities for the types in Wave 3, dependent on their type in Wave 2. We then assign to each individual the type with the highest predicted probability and list our observations in a so-called confusion matrix. Model 1 in Table 3 shows the results. The model predicts the type Conditional Cooperator for all subjects who were either Conditional Cooperators or Free Riders in Wave 2 (54 subjects), the remaining 14 subjects are predicted to remain in the category Other. In Fig. 2 it is easy to see why this is the case. Between Wave 1 and Wave 2 the largest number of both Free Riders and Conditional Cooperators moves to the category Conditional Cooperator. The nine numbers in the center of the table for Model 1 demonstrate the accuracy of the prediction. Among the 54 subjects who are predicted to be Conditional Cooperators, 33 turn out to be of that type in Wave 3, 11 are Free Riders and 10 are Others. Bold numbers on the main diagonal are the observations for which the prediction is correct. The predictive success of the model is simply the number of correct predictions in relation to all observations in Wave 3 (68). The rightmost column shows the accuracy of Model 1, which is in this case 58.8 percent.

In Model 2 we use the Big-Five personality dimensions to estimate a multinomial logit model with data from Wave 1 and Wave 2. This model predicts only the types Conditional Cooperator and Free Rider. This reflects the results discussed in the previous section that personality measures are informative for the distinction between Free Riders and Conditional Cooperators, but not when it comes to the category Others. Model 2 does not take the transition probabilities into account. The personality measures allow us to predict the correct type in 66.2 percent of the cases in Wave 3.

In Model 3 we combine the personality measures with the lagged type of a subject. This model predicts the occurrence of all three types and achieves an accuracy of 67.6 percent. Finally, in Model 4 we introduce interaction terms between the personality measures and the lagged type. This allows accounting for the fact that personality and the probability of changing type between two points in time are likely to be related. To keep the number of explanatory variables low we only interact the most important personality dimension (Agreeableness) with the lagged type. Model 4 improves the prediction relative to Model 3 and achieves an accuracy of 70.6 percent.

[19] If p_i is the predicted probability of type i (i = 1, . . ., n) and q_i is the true probability, then the prediction is correct with probability $\Sigma_{p_i q_i}$. If all p_i are identical then the sum reduces to $1/n$. Thus for three types the hit rate is one third, independent of the true probabilities.

Accuracy might not be the only performance measure we are interested in. An alternative approach is to ask which model performs best in detecting a specific type, or, in other words, avoids type I errors (false positives). If we look at the type Conditional Cooperator, clearly Model 2 has the lowest type I error rate. All 38 Conditional Cooperators observed in Wave 3 are predicted to be of that type. For the type Free Rider all models have much higher type I errors. Model 4 performs best, correctly identifying 8 out of the 13 Free Riders we observe in Wave 3. In all models type I errors in predicting the type Other are high. Finally, we could also look at type II errors, i.e., the case where, e.g., a player is predicted to be a Conditional Cooperator when in fact he is not (false negatives). The rate of type II errors is large in Models 1 and 2, where 70 and 80 percent of the non-Conditional Cooperators are predicted to be Conditional Cooperators. Even Models 3 and 4 have a type II error rate of 43 percent. On the other hand, type II error rates are very low for Free Riders in all four models. We conclude that, despite the fact that a substantial fraction of the subjects change their type between waves, information about subjects' past behavior in combination with personality measures is highly informative when we aim at predicting individual types.

4. Conclusions

Economic theory almost always implicitly assumes that preferences are stable. Yet, empirical evidence on the stability of preferences is very limited. Furthermore, economists routinely classify preferences into different types without establishing the underpinnings of preference heterogeneity. The goal of this study was to investigate the stability of cooperation preferences over time and to explore the psychological foundations of preference heterogeneity and preference stability. The main contribution of our study is threefold.

First, we report high aggregate stability of cooperation preferences over time. This finding complements Blanco et al. (2011) who report high aggregate stability of social preferences across different games. Taken together, these results suggest that theories of social preferences can "provide a system of generalizations that can be used to make correct predictions about the consequences of any change in circumstances", as stipulated by Friedman (1953, p. 4). Our finding of high aggregate stability is also interesting, as it suggests that the heterogeneity of other-regarding preferences reported in previous studies is temporally stable within the same subject pool. In line with earlier research, we find that the most important preference types are Conditional Cooperators and Free Riders, which has important theory and policy consequences (see Gächter, 2007 for a comprehensive discussion).

Second, we provide evidence on the stability of individual preferences over time. Individual-level preference stability is an important prerequisite for signaling games or reputation mechanisms. Learning another players' type through signaling makes only sense if there is such a thing as a stable type (see e.g., the models by Ellingsen and Johannesson (2008) or Levine (1998)). We find that individual-level stability of cooperation preferences over time is lower than aggregate stability. Still, individual stability is high enough so that knowledge about an individual's past behavior offers substantial informational content when it comes to predicting future behavior. Our results also highlight another interesting aspect. While previous work has focused on one dimension of preference heterogeneity, i.e., heterogeneity in terms of preference types, our findings suggest there is a second dimension researchers should pay attention to, i.e., heterogeneity in terms of preference stability. Some individuals seem to have stable preferences, while others do not. However, given the small number of studies on the within-subject stability of cooperation preferences, more research is certainly needed for a proper understanding of this phenomenon.

Third, we provide evidence on the psychological foundations of cooperation preferences. Our results suggest that the Free Riders and Conditional Cooperators identified in our experiments are subjects with different personalities, as measured by the Big-Five dimension 'Agreeableness'. The relation between personality and cooperation preferences is particularly strong for consistent Conditional Cooperators and consistent Free Riders. Information about subjects' personality traits improve the accuracy of our predictions for Wave 3 types beyond what was possible relying only on behavioral data.

In conclusion, our results indicate that there is heterogeneity of individual behavior in terms of both preference types and preference stability, which can be linked to individual differences in personality. Applying an individual differences approach to investigate the importance of dispositional factors may accordingly prove useful for our understanding of the heterogeneity and stability of cooperation preferences. We believe that future theoretical and empirical research on social preferences would benefit substantially from incorporating personality and attitude variables. We therefore call for more interdisciplinary research that spans the behavioral economics and personality psychology fields to cross-fertilize insights from both disciplines.

Acknowledgements

We are grateful for helpful comments from the editors, two anonymous referees, Therese Faessler, Simon Gächter, Jonathan Schulz, and the participants of the Thurgau Experimental Economics Meeting 2010 in Kreuzlingen, Switzerland.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2011.10.006.

References

Almlund, M., Duckworth, A.L., Heckman, J.J., Kautz, T. Personality psychology and economics. In: Hanushek, E.A., Machin, S., Wößmann, L. (Eds.), Handbook of the Economics of Education, vol. 4. Elsevier, Amsterdam, forthcoming.

Ameriks, J., Caplin, A., Leahy, J., Tyler, T., 2007. Measuring self-control problems. American Economic Review 97, 966–972.

Andersen, S., Harrison, G.W., Lau, M.I., Rutström, E., 2008. Lost in state space: are preferences stable? International Economic Review 49, 1091–1112.

Andreoni, J., Miller, J., 2002. Giving according to Garp: an experimental test of the consistency of preferences for altruism. Econometrica 70, 737–753.

Ashton, M.C., Paunonen, S.V., Helmes, E., Jackson, D.N., 1998. Kin altruism, reciprocal altruism, and the Big Five personality factors. Evolution and Human Behavior 19, 243–255.

Bartling, B., Fehr, E., Marechal, M.A., Schunk, D., 2009. Egalitarianism and competitiveness. American Economic Review 99, 93–98.

Becker, G.S., 1976. The Economic Approach to Human Behavior. University of Chicago Press, Chicago.

Ben-Ner, A., Kong, F., Putterman, L., 2004a. Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving. Journal of Economic Psychology 25, 581–589.

Ben-Ner, A., Kramer, A., 2011. Personality and altruism in the dictator game: relationship to giving to kin, collaborators, competitors, and neutrals. Personality and Individual Differences 51, 216–221.

Ben-Ner, A., Putterman, L., Kong, F., Magan, D., 2004b. Reciprocity in a two-part dictator game. Journal of Economic Behavior & Organization 53, 333–352.

Blanco, M., Engelmann, D., Normann, H.-T., 2011. A within-subject analysis of other-regarding preferences. Games & Economic Behavior 72, 321–338.

Boone, C., De Brabander, B., Carree, M., de Jong, G., van Olffen, W., van Witteloostuijn, A., 2002. Locus of control and learning to cooperate in a Prisoner's Dilemma game. Personality and Individual Differences 32, 929–946.

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

Brandstätter, H., Königstein, M., 2001. Personality influences on ultimatum bargaining decisions. European Journal of Personality 15 (S1), S53–S70.

Brandts, J., Charness, G., 2011. The strategy versus the direct–response method: a first survey of experimental comparisons. Experimental Economics 14, 375–398.

Brosig, J., Riechmann, T., Weimann, J., 2007. Selfish in the end? An investigation of consistency and stability of individual behavior. MPRA Paper 2035, University Library of Munich.

Burlando, R.M., Guala, F., 2005. Heterogeneous agents in public goods experiments. Experimental Economics 8, 35–54.

Buss, D.M., 1991. Evolutionary personality psychology. Annual Review of Psychology 42, 459–491.

Chmielewski, M., Watson, D., 2009. What is being assessed and why it matters: the impact of transient error on trait research. Journal of Personality and Social Psychology 97, 186–202.

Conley, J.J., 1985. Longitudinal stability of personality traits: a multitrait–multimethod–multioccasion analysis. Journal of Personality and Social Psychology 49, 1266–1282.

Costa, P.T., McCrae, R.R., 1992. Revised NEO Personality Inventory (NEO PI-R) and NEO Five-factor Inventory (NEO FFI) Professional Manual. Psychological Assessment Resources, Odessa, FL.

Cox, J.C., 2004. How to identify trust and reciprocity. Games & Economic Behavior 46, 260–281.

De Oliveira, A.C.M, Croson R.T.A., Eckel, C., 2009. Are preferences stable across domains? An experimental investigation of social preferences in the field. CBEES Working Paper # 2008-3.

Donnellan, M.B., Oswald, F.L., Baird, B.M., Lucas, R.E., 2006. The mini IPIP scales: tiny yet effective measures of the Big Five factors of personality. Psychological Assessment 18, 192–203.

Dufwenberg, M., Kirchsteiger, G., 2004. A theory of sequential reciprocity. Games & Economic Behavior 47, 268–298.

Ehrhart, M.G., Ehrhart, K.H., Roesch, S.C., Chung-Herrera, B.G., Nadler, K., Bradshaw, K., 2009. Testing the latent factor structure and construct validity of the Ten-Item Personality Inventory. Personality and Individual Differences 47, 900–905.

Ellingsen, T., Johannesson, M., 2008. Pride and prejudice: the human side of incentive theory. American Economic Review 98, 990–1008.

Falk, A., Fischbacher, U., 2006. A theory of reciprocity. Games & Economic Behavior 54, 293–316.

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

Ferguson, E., Heckman, J.J., Corr, P., 2011. Personality and economics: overview and proposed framework. Personality and Individual Differences 51, 201–209.

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10, 171–178.

Fischbacher, U., Gächter, S., 2009. The behavioral validity of the strategy method in public good experiments. CeDEx Discussion Paper No. 2009-25, University of Nottingham.

Fischbacher, U., Gächter, S., 2010. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. American Economic Review 100, 541–556.

Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Economic Letters 71, 397–404.

Fisman, R., Kariv, S., Markovits, D., 2007. Individual preferences for giving. American Economic Review 97, 1858–1876.

Fleming, P., Zizzo, D.J., 2011. Social desirability, approval and public good contribution. Personality and Individual Differences 51, 258–262.

Friedman, M., 1953. The methodology of positive economics. In: Friedman, M. (Ed.), Essays in Positive Economics. University of Chicago Press, Chicago, pp. 3–43.

Gächter, S., 2007. Conditional cooperation. Behavioral regularities from the lab and the field and their policy implications. In: Frey, B.S., Stutzer, A. (Eds.), Economics and Psychology. A Promising New Cross-disciplinary Field. MIT Press, Cambridge, pp. 19–50.

Gächter, S., Herrmann, B., Thöni, C., 2004. Trust, voluntary cooperation, and socio-economic background: survey and experimental evidence. Journal of Economic Behavior & Organization 55, 505–531.

Gächter, S., Herrmann, B., Thöni, C., 2010. Culture and cooperation. Philosophical Transactions of the Royal Society B 365, 2651–2661.

Gächter, S., Thöni, C., 2005. Social learning and voluntary cooperation among like-minded people. Journal of the European Economic Association 3, 303–314.

Goldberg, L.R., 1993. The structure of phenotypic personality traits. American Psychologist 48, 26–34.

Gosling, S.D., Rentfrow, P.J., Swann, W.B., 2003. A very brief measure of the Big-Five personality domains. Journal of Research in Personality 37, 504–528.

Graziano, W.G., Eisenberg, N.H., 1997. Agreeableness: a dimension of personality. In: Hogan, R., Johnson, J., Briggs, S. (Eds.), Handbook of Personality Psychology. Academic Press, London, pp. 795–824.

Gunnthorsdottir, A., McCabe, K., Smith, V., 2002. Using the Machiavellianism instrument to predict trustworthiness in a bargaining game. Journal of Economic Psychology 23, 49–66.

Herrmann, B., Thöni, C., 2009. Measuring conditional cooperation: a replication study in Russia. Experimental Economics 12, 87–92.

Hirsh, J., Peterson, J.B., 2009. Extraversion, neuroticism, and the prisoner's dilemma. Personality and Individual Differences 46, 254–256.

Horowitz, J.K., 1992. A test of intertemporal consistency. Journal of Economic Behavior & Organization 17, 171–182.

John, O.P., Srivastava, S., 1999. The Big Five trait taxonomy: history, measurement, and theoretical perspectives. In: Pervin, L.A., John, O.P. (Eds.), Handbook of Personality: Theory and Research. Guilford Press, New York, pp. 102–138.

Kocher, M.G., Cherry, T., Kroll, S., Netzer, R.J., Sutter, M., 2008. Conditional cooperation on three continents. Economics Letters 101, 175–178.

Koole, S.L., Jager, W., van den Berg, A.E., Vlek, C.A.J., Hofstee, W.K.B., 2001. On the social nature of personality: effects of extraversion, agreeableness, and feedback about collective resource use on cooperation in a resource dilemma. Personality and Social Psychology Bulletin 27, 289–301.

Kurzban, R., Houser, D., 2001. Individual differences in cooperation in a circular public goods game. European Journal of Personality 15, S37–S52.

Kurzban, R., Houser, D., 2005. Experiments investigating cooperative types in humans: a complement to evolutionary theory and simulations. Proceedings of National Academy of Science of the United States of America 102, 1803–1807.

Levine, D.K., 1998. Modeling altruism and spitefulness in experiments. Review of Economic Dynamics 1, 593–622.
Lo, A., Repin, D.V., Steenbarger, B.N., 2005. Fear and greed in financial markets: a clinical study of day traders. American Economic Review 95, 352–359.
McCrae, R.R., Costa, P.T., 1990. Personality in Adulthood. Guilford Press, New York.
Meier, S., Sprenger, C., 2010. Stability of time preferences. IZA Discussion Paper No. 4756.
Meyer, H.-D., 1992. Norms and self-interest in ultimatum bargaining: the prince's prudence. Journal of Economic Psychology 13, 215–232.
Muller, L., Sefton, M., Steinberg, R., Vesterlund, L., 2008. Strategic behavior and learning in repeated voluntary contribution experiments. Journal of Economic Behavior & Organization 67, 782–793.
Nunnally, J.C., Bernstein, I.H., 1994. Psychometric Theory, 3rd edition. McGraw-Hill, New York.
Olver, J.M., Mooradian, T.A., 2003. Personality traits and personal values: a conceptual and empirical integration. Personality and Individual Differences 35, 109–125.
Parks, L., Guay, R.P., 2009. Personality, values, and motivation. Personality and Individual Differences 47, 675–684.
Perugini, M., Tan, J.H., Zizzo, D.J., 2010. Which is the more predictable gender? Public good contribution and personality. Economic Issues 15, 83–111.
Pothos, E.M., Perry, G., Corr, P.J., Matthew, M.R., Busemeyer, J.R., 2011. Understanding cooperation in the Prisoner's Dilemma game. Personality and Individual Differences 51, 210–215.
Rapoport, A., Suleiman, R., 1993. Incremental contribution in step-level public goods games with asymmetric players. Organizational Behavior and Human Decision Processes 55, 171–194.
Skatova, A., Ferguson, E., 2011. What makes people cooperate? Individual differences in BAS/BIS predict strategic reciprocation in a public goods game. Personality and Individual Differences 51, 237–241.
Thöni, C., Tyran, J., Wengström, E., 2009. Microfoundations of social capital. Discussion Paper No. 09-24, University of Copenhagen.
Volk, S., Thöni, C., Ruigrok, W., 2011. Personality, personal values and cooperation preferences in public goods games: a longitudinal study. Personality and Individual Differences 50, 810–815.
Zeisberger, S., Vrecko, D., Langer T. Measuring the time stability of Prospect Theory preferences. Theory and Decision, in press.