Personality accounts for stable preferences and expectations across a range of simple games

Sean Brocklebank a,*, Gary J. Lewis b, Timothy C. Bates b

a School of Economics, University of Edinburgh, United Kingdom
b Department of Psychology, University of Edinburgh, United Kingdom

ABSTRACT

Behaviour on even simple experimental games shows considerable individual differences, but previous attempts to link these preferences to stable personality traits have had mixed results. Here we address three limitations of earlier studies, namely: (1) uncertainties concerning the reliability of preferences; (2) use of personality instruments with limited cross-study comparability; and (3) confounds where more than one psychological motive can lead to a particular choice. Sixty-seven participants completed 12 distinct real-money games twice over a two-week interval along with 6 measures concerning their expectations about other players' choices. Personality was measured using the full NEO-PI-R. Choices were highly stable across time (r = .84). Moreover, choices on the 12 games and 6 expectations reflected a single underlying dimension of "prosocial orientation", measuring concern for the payoffs received by other players. Scores on the prosocial orientation dimension were related to personality, with openness, (low) neuroticism, and (low) extraversion retained as significant predictors.

1. Introduction

According to Binmore (2007), "A game is being played whenever people have anything to do with each other" (p. 3). As used in research, games are generally run in 2-player laboratory settings. In a typical dictator game, a dictator is endowed with a fund from which they must choose some amount (from 0% to 100%) to give to the recipient. The variable under study is the percent offered by the dictator. Multiple variations of such games have been developed: For instance, in the ultimatum game the recipient can choose to accept or reject the offer (in the case of rejection neither player receives anything).

Research has revealed considerable individual differences in social preferences on these simple games (defined here as preferences over the distribution of resources between individuals; Camerer, 2003). One candidate in explaining these important differences is personality, where prima-facie associations, such as links of agreeableness to empathy and cooperation (Jensen-Campbell & Graziano, 2001), suggest associations with benevolent social preferences. However, studies testing these associations have reported mixed results (e.g. Kurzban & Houser, 2001).

Here we present a study of the influences of personality on social preferences taking into account three possible limiting factors in previous research, namely: (1) the existence of inherent confounds within certain games used in prior research, such that identical behaviours can reflect distinct underlying motivations; (2) the limited comparability of personality instruments used in previous research; and (3) the possible low-stability of social preferences. Next we briefly introduce previous work examining relationships between personality and social preferences, before describing limitations in previous research in more detail, and, finally, presenting a study that addresses these limitations.

1.1. Individual differences in social preferences

It has long been argued that individual differences are likely to play at best a trivial role in determining social preferences (Pruitt & Kimmel, 1977), though personality has been linked to retaliation (Skarlicki, Folger, & Tesluk, 1999) and to preferences over allocations (Schmitt, Neumann, & Montada, 1995), both of which are intimately related to social preferences. Recent studies, however, have begun to explore trait dispositions underlying variation in economic behaviour generally (Borghans, Duckworth, Heckman, & ter Weel, 2008) and in game behaviour specifically. For example, Hirsh and Peterson (2009) found that the withdrawal aspect of neuroticism (tapping fear and insecurity) and the enthusiasm aspect of extraversion (tapping positive affect and sociability) from
the Big Five aspect scale (DeYoung, Quilty, & Peterson, 2007) independently predicted a greater likelihood of cooperation in a prisoner's dilemma game ($\beta = -.14, -.12$, respectively). By contrast, Lönqvist, Verkasalo, and Walkowitz (2011) found that low neuroticism and high Openness to Experience predicted more cooperative transfers. Using the dictator game paradigm, Ben-Ner, Kong, and Putterman (2004) reported significant associations between agreeableness and (low) extraversion and the sum offered by the dictator. Finally, Kurzan and Houser (2001) reported non-significant associations between Big Five personality traits and social preferences. Further studies have examined variation in social preferences using personality frameworks other than the five-factor model. For example, Boone, De Brabander, and van Witteloostuijn (1999) observed that the personality traits locus of control, self-monitoring, and sensation seeking had significant associations ($r = .28-.44$) with levels of cooperative behaviour in a prisoners’ dilemma game. Scheres and Sanfey (2006) observed significant associations between BAS-Drive and BAS-Reward and (low) offers in a dictator game. And Swope, Cadigan, Schmitt, and Shupp (2008) reported no significant effects of the Myers–Briggs Type Indicator on social preferences.

1.2. Limitations in previous research

There are a number of possible explanations for the mixed results described above. Firstly, much research has focused on just one or two experimental games, such as the dictator and ultimatum games; however, important confounds have been identified in these games which render choices ambiguous as to underlying motivations or preferences (Charness & Rabin, 2002). For example, rejection of a low offer in the ultimatum game can reflect difference aversion or retaliation. These distinct motives, in turn, confound potential underlying personality traits, such as neuroticism and agreeableness, respectively. Likewise, in the prisoners’ dilemma, a choice to defect can reflect aversion to differential outcomes, aversion to risk, or a self-regarding preference. These confounds can be mitigated by exploring a range of payoff pairings, varying absolute and relative payoff differences, as well as allowing multi-stage games (Charness & Rabin, 2002). Finally, and importantly, choices reflect expectations about the other player in addition to personal preferences. An example would be the expectation (or fear) that the other player will defect. Because of these confounds in single games, personality is likely to have apparently divergent or null associations to preferences on different games because of the distinct ways in which each game might trigger personality-related preferences.

Secondly, the various personality instruments used in studies associating social preferences and personality have made it difficult to compare results and uncover personality-preference links. For example, Swope et al. (2008) used the Myers–Briggs Type Indicator (which does not tap neuroticism; McCrae & Costa, 1989), and Boone et al. (1999) used an assortment of scales: locus of control, self-monitoring, type-A behaviour, and sensation-seeking. While each of these measures may tap specific traits of relevance to social preferences, the core five-factor model has demonstrated broader coverage of stable human behaviour than any other measurement instrument (e.g. Costa & McCrae, 1992; Goldberg, 1993), and so provides a more comprehensive tool by which to understand putative trait influences on social preferences.

Finally, and most straightforwardly, research has seldom addressed the reliability of social preferences. As recently noted by Ferguson, Heckman, and Corr (2011), the stability of economic preferences still needs to be established. Although we do not think that this is the likely explanation for the mixed results, if reliability in choice behaviour is low (e.g. because participants choose randomly), this would explain both the high variability typically seen in games and the inconsistency of measured relationships with stable personality traits in previous research, as noted above.

1.3. The current study

To address these limitations, in the present study we measured social preference with Charness and Rabin’s (2002) set of dictator games (described in more detail below) twice over a two-week interval, and utilised the full-spectrum NEO-PI-R (Costa & McCrae, 1992) in order to gain a comprehensive assessment of personality. The current study used a large set of games which are well-established in the experimental economics literature (Charness & Rabin, 2002). This mixture of games allows us to aggregate over many more choices than are commonly elicited from subjects and thus to eliminate common confounds between Pareto-damaging behaviour (behaviour that makes at least one person worse off without making anyone better off, in monetary terms), retaliation, and inequality reduction. These games also tap into the two primary factors which economic theorists have identified as critical for explaining social preferences: How much the other participants receive (comparison-based preferences; people will be less kind towards those who have more than themselves), and the perceived intentions of the other participants (intention-based preferences; people will be less kind towards those who have shown bad intentions). These factors have been separately identified by Fehr and Schmidt (1999), Bolton and Ockenfels (2000), and Charness and Rabin (2002; see also Daruvala, 2010, for a review), but have so far only been discussed in terms of their influence on average behaviour: The factor structure of these games has not yet (to our knowledge) been examined.

1.4. Predictions

With regard to social preferences, we were agnostic about the underlying factor structure on account of the scarcity of individual differences work in this field to guide predictions. In addition to assessing the reliability of social preferences and examining the consistency of these preferences across a range of 18 games, we made several predictions relating personality to social preferences. Concerning comparison-based preferences, we hypothesised that agreeableness would be positively associated with choices reflecting concern for the welfare of others, as well as positive expectations of others’ choices, on account of demonstrated links with empathy and trust (Jensen-Campbell & Graziano, 2001). Similarly, we predicted that neuroticism would associate negatively with concern for welfare of others, and expectations of others’ choosing selfishly, due to the contribution of facets such as hostility (Costa & McCrae, 1992). Finally, we hypothesised that openness would predict benevolent social preferences, and positive expectations of others’ choices, on account of relationships of openness to values of fairness and harm reduction (Lewis & Bates, 2011a). Regarding predictions concerning personality associations with intention-based preferences, we hypothesised that neuroticism and extraversion, with links to revengeful thoughts following a transgression (Malby et al., 2008) and dominance behaviours (Nettle, 2005), respectively, would predict less concern for the welfare of others following a selfish choice.

2. Method

2.1. Participants

Seventy-five participants were recruited from an undergraduate participation pool: participants received partial course credit for attending as well as a financial remuneration based on choices...
made in the experimental tasks. Of the initial 75 participants, 71
returned for the second session. Additionally, 4 participants’ data
were lost due to a data storage failure. Of the 67 remaining partic-
ipants, 54 were female (mean age = 19; modal age = 18;
range = 17–50; SD = 3.9 years).

2.2. Measures

2.2.1. Personality

Five-factor model (FFM) personality traits were measured using
the 240-item NEO-PI-R (Costa & McCrae, 1992). Participants com-
pleted the inventory at individual computer terminals.

2.2.2. Games

A set of six dictator and six response games (described below)
were taken from Charness and Rabin (2002) and are named accord-
ing to their convention (with the exception of Ed 128, which is de-

erived from Berk 28 but was not in the original set of games). As an
example of comparison-based preferences, in the game known as
Berk 23 (see Fig. 1), Player B chooses between an outcome in which
Player A player gets £8 and Player B gets £2, versus an outcome in
which each receives £0. As an example of an intention-based pref-

erence, in the game known as Berk 22 (see Fig. 2), Player A can
choose £3.75 for themselves and £10 for Player B, or let Player B
choose between £4 for each player or £2.50 for Player A and
£3.50 for themselves. Here, if Player A ‘enters’ the game and allows
Player B to make the choice, Player A deprives Player B of a guar-
anteed £10. Of course, Player B may now choose the lower payoff
for themselves (£3.50 rather than £4) in order to punish Player A
(Player A would then receive £2.50 rather than £4). Participants
played all response games both as Player B and as Player A.

Games are listed in Table 1 corresponding to the (fixed) order
that they were played by the participants. Games are presented
so that the prosocial choice for Player B is always on the left
(although the games were counter-balanced when presented to the
participants), with the exception of Berk 26, in which the total
payoff is identical for both choices available to Player B.

In order to explore the role of players’ expectations about the
behaviour of others on their own choices, participants were asked
to estimate the percentage of all other participants who would
make the prosocial choice when acting as Player B in the relevant
games. This was taken after they had made their choice in the role
of Player A in the response games. Participants were informed that
there would be a £10 prize for the participant with the most accu-
rate estimates of other players’ behaviour.

2.3. Procedure

Participants were tested individually in separate experimental
cubicles. Participants were informed both when they signed up
for the experiment and again at the beginning of their first session
that they would be required to return for the second half of the
experiment in two weeks in order to obtain both course credit
and monetary payment. Participants were paid at the end of the
second session based on their rewards in one task from each of
the two sessions chosen at random. All of this was common
knowledge.

Participants first played all six dictator games before playing all
six response games as the second player and finally playing the
same six response games as the first player (see Table 1). Particip-
ants were told that payoffs would be based on converting win-
ings from one game at random in each session into British
pence (i.e. 750 = £7.50). Participants were not told of their part-
ners’ choices until the end of the second session, when they were
paid. The NEO-PI-R was administered in two blocks: one at the
end of the first session and one at the end of the second session.

3. Results

The proportions of choices made by participants for each game,
and the expectations of other players’ behaviour, are summarised
in Table 1. Descriptive statistics for the personality traits are de-
tailed in Table 2.

3.1. Factor structure of social preferences

To our knowledge, the factor structure of social preferences has
not previously been determined: Accordingly, we performed
exploratory factor analysis on the game and expectation data. Be-
cause Charness and Rabin have proposed three parameters under-
lying social preferences we conducted FA to see whether behaviour
our 18 games would reflect this as a 3-factor structure. Eigenvalues
for the varimax rotated solution were (7.6, 4.3, 2.4, 1.8, 0.9, ...
...), sug-

... Fig. 1. Sample screenshot of the Berk 23 game as presented to participants.
loading on the first factor) and the third factor was not interpretable in terms of the Charness and Rabin structure. Jointly this suggested extracting one factor. The one-factor solution loaded positively on all items and accounted for 58% of the variance.

We next aggregated game preferences into a single construct which we term the Prosocial Orientation Scale (POS). For simplicity and to avoid capitalising on chance factor loadings, POS scores were calculated as a sum of choices rather than from the factor loadings. POS scores were derived as follows: one point was awarded for each game on which the player chose the option where the total payoff (for both players) was largest, i.e. the prosocial choice, and, for expectations a score between 0 and 1 was awarded representing the expectation of prosociality in other players. The POS demonstrated excellent internal-consistency (Cronbach’s $\alpha = .94$) and high test–retest reliability (across the two week interval; $r = .84$, $p < .0001$).

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Out</th>
<th>Enter</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dictator games</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berk 17</td>
<td>.16</td>
<td>.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berk 23</td>
<td>.67</td>
<td>.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berk 29</td>
<td>.38</td>
<td>.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berk 15</td>
<td>.50</td>
<td>.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berk 26</td>
<td>.35</td>
<td>.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ed 128</td>
<td>.72</td>
<td>.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Response games</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berk 13</td>
<td>.86</td>
<td>.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berk 30</td>
<td>.49</td>
<td>.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berk 31</td>
<td>.78</td>
<td>.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berk 19</td>
<td>.75</td>
<td>.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berk 22</td>
<td>.31</td>
<td>.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berk 28</td>
<td>.31</td>
<td>.69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note**: Numbers in parentheses show (Player A, Player B) payoffs in British pence; Out = The proportion of Player As who opted to stay ‘out’ of the game by choosing the available payoffs and thus depriving Player B of a choice; Enter = The proportion of Player As who ‘entered’ the game, thus allowing Player B to make a choice between the available payoffs; Left = The proportion of Player Bs choosing the left-hand payoff bundle; Right = the corresponding proportion of Bs who chose the right-hand bundle. Expectations of B = Player A’s expectations of the average B choice.

### Table 2

Descriptive statistics for the five personality traits.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroticism</td>
<td>57.64</td>
<td>9.87</td>
</tr>
<tr>
<td>Extraversion</td>
<td>55.12</td>
<td>8.61</td>
</tr>
<tr>
<td>Openness</td>
<td>56.03</td>
<td>7.02</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>47.20</td>
<td>8.52</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>41.99</td>
<td>8.67</td>
</tr>
</tbody>
</table>

**Note**: Means are $t$-scores derived from manual norms.

3.2. Personality as a predictor of prosocial orientation

Having established that social preferences demonstrated high reliability and stability, we next examined the relationship of FFM traits to social preferences using linear modelling (multiple regression) with POS scores as the dependent variable and entering
each of the FFM domains (as well as age and sex) as independent variables. This model accounted for 25.5% of variance in POS scores, with neuroticism ($\beta = -0.33, p = 0.02$), extraversion ($\beta = -0.32, p = 0.02$), and openness to experience ($\beta = 0.41, p = 0.003$) being significant predictors (see Table 3). Agreeableness, conscientiousness, sex, and age were not significant predictors and removing these variables did not significantly alter model fit ($F(1, 46, p = 0.23$) or the parameter estimates of the significant predictors (though we did not have high power to detect sex effects in the current study). Nor did the inclusion of facets improve the fit: when we added either the 18 facets from neuroticism, extraversion and openness (the significant predictors) or all 30 facets, the joint significance tests could not reject the null hypothesis of zero predictive power ($F = 1.04, p = 0.44$ and $F = 1.46, p = 0.15$, respectively). Pairwise interaction terms for all personality factors were non-significant. Similarly, moving from measuring POS based on a simple sum (as above) to a factor score revealed the same pattern of significance and made no meaningful difference to coefficient magnitudes: coefficients on neuroticism, extraversion, and openness went from $-0.33, -0.32$ and $0.41$ in the summation model to $-0.38, -0.34$ and $0.34$ in the factor model, respectively.

<table>
<thead>
<tr>
<th>Table 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personality predictors of the Prosocial Orientation Scale (standardised coefficients are presented) for Model 1 and 2.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
</tr>
<tr>
<td>Extraversion</td>
</tr>
<tr>
<td>Openness</td>
</tr>
<tr>
<td>Agreeableness</td>
</tr>
<tr>
<td>Conscientiousness</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Sex (male = 0)</td>
</tr>
<tr>
<td>Multiple-R²</td>
</tr>
</tbody>
</table>

Note. $N = 67$. 

$^{*} p < 0.05$. 

$^{**} p < 0.01$. 

In conclusion, the results suggest that differences in behaviour on simple games are stable, that they reflect a general preference for prosocial outcomes, and that they have a significant link to personality traits of extraversion, neuroticism, and openness. Future work seeking to identify trait associations with social preferences is recommended to place less emphasis on confounded games, such as the prisoners’ dilemma and the ultimatum game, and instead make greater use of games that avoid such drawbacks. Extensions to our preliminary investigation of the psychometric structure of social preferences will also be valuable.

**References**
