Prosocial phenotype predicts political views on hierarchy and redistribution eighteen

months later

Scott Claessens[1*], Chris G Sibley[1], Ananish Chaudhuri[2,3], & Quentin D Atkinson[1,4]

[1] School of Psychology, University of Auckland, Auckland, New Zealand

[2] Department of Economics, University of Auckland, Auckland, New Zealand

[3] CESifo, Munich, Germany

[4] Max Planck Institute for the Science of Human History, Jena, Germany

* Correspondence concerning this article should be addressed to Scott Claessens, Level 2, Building 302, 23 Symonds Street, Auckland, New Zealand. E-mail: scott.claessens@gmail.com

Word count = 4983 words

Author biographies:

- **Scott Claessens** is a Research Fellow in the School of Psychology at the University of Auckland. He is interested in the evolution of cooperation, culture, and political ideology.

- **Chris G Sibley** is a Professor in the School of Psychology at the University of Auckland. He is the Principal Investigator for the New Zealand Attitudes and Values Study.

- **Ananish Chaudhuri** is a Professor of Experimental Economics in the Department of Economics at the University of Auckland. He has a broad research agenda in behavioural economics, including a significant amount of cross-disciplinary work.

- **Quentin D Atkinson** is a Professor in the School of Psychology at the University of Auckland and co-Director of the University of Auckland Behavioural Insights Exchange (UoABIX). He is interested

in the origins of linguistic diversity, the function of religion, the psychology of climate change, and how evolved cognitive biases shape our social behaviour.

Abstract

Cross-sectional research has identified robust correlations between prosocial behaviour in economic games and political views, but this research is limited in its ability to draw causal inferences. Here, we conducted a longitudinal cross-lagged panel study of prosociality and political views with a New Zealand sample (n = 631). Across two waves separated by eighteen months, we measured self-reported political views and employed a battery of economic games to estimate people's general preferences for prosociality. We found that this "prosocial phenotype" predicted future variation in some of our measures of political views, including Social Dominance Orientation and support for income redistribution. Income attribution beliefs and political party support were not consistently related to the prosocial phenotype over time. None of these variables predicted future variation in the prosocial phenotype. These results suggest that prosocial predispositions may play a causal role in the expression of certain political views.

Prosocial phenotype predicts political views on hierarchy and redistribution eighteen months later

Introduction

People vary in their political voting patterns, policy views, and party support. In part, this variation is thought to be due to individual differences in political ideology, defined as "a set of stable, interrelated beliefs and attitudes that organise views on political and social issues" (Jost, Federico, & Napier, 2009).

It has long been assumed that one's ideology is primarily acquired through partisanship that is learned from one's parents and peers or soaked up through mass media (Campbell, Converse, Miller, & Stokes, 1960; Zaller, 1992). This assumption continues today, with researchers arguing that differing political views are due mainly to partisan divides rather than any enduring psychological or dispositional differences (Brandt, Reyna, Chambers, Crawford, & Wetherell, 2014). However, this claim is at odds with research highlighting the importance of dispositional variables in the expression of political ideology. This body of work has revealed that, compared to liberals, conservatives tend to be more sensitive to threats and uncertainty and express a greater need for closure (Jost, Glaser, Kruglanski, & Sulloway, 2003). Different variables have also been linked to different measures of political ideology. Social conservatives and people higher in Right Wing Authoritarianism (RWA; an ideological measure of norm-adherence and group conformity) (Altemeyer, 1981) tend to be more sensitive to disgusting stimuli (Inbar, Pizarro, Iyer, & Haidt, 2012), while economic conservatives and people higher in Social Dominance Orientation (SDO; an ideological measure of support for hierarchy and dominance) (Pratto, Sidanius, Stallworth, & Malle, 1994) tend to exhibit higher levels of physical dominance and formidability (Petersen & Laustsen, 2019).

Prosocial preferences have also been shown to vary across the ideological landscape. Research has found a negative relationship between SDO and Social Value Orientation

(SVO), an other-regarding preference for allocating resources to others (Chirumbolo, Leone, & Desimoni, 2016; Haesevoets, Reinders Folmer, & Van Hiel, 2015). Research has also revealed robust correlations between political views and prosocial behaviour in economic games. Economic games (*i.e.*, social decision-making tasks that involve real money) are tools that elicit prosocial preferences such as trust, cooperation, and reciprocity while avoiding the desirability issues that plague self-report methods (Pisor, Gervais, Purzycki, & Ross, 2020). Studies using economic games have shown that people higher in SDO tend to share less money in social dilemma games that pit self-interest against collective-interest (Haesevoets, Reinders Folmer, Bostyn, & Van Hiel, 2018; Haesevoets et al., 2015; Halali, Dorfman, Jun, & Halevy, 2018). A recent meta-analysis of data from over 3,000 participants found a reliable negative correlation between SDO and prosocial behaviour in economic games, albeit with a small effect size (Thielmann, Spadaro, & Balliet, 2020).

Extending this work, Claessens, Sibley, Chaudhuri, and Atkinson (2023) found that SDO was negatively correlated with a general behavioural disposition for prosociality that is uncovered via a battery of economic games (Peysakhovich, Nowak, & Rand, 2014). These economic games included the Trust Game, Dictator Game, and Public Goods Game. Although different motivations underlie behaviour in these games, such as cooperation, altruism, and fairness, prosocial behaviour in all of these games reflects a concern for the welfare of others when there is a possibility for exploitation (Thielmann, Hilbig, & Zettler, 2022). The common factor underlying prosocial behaviour in these economic games, referred to here as the "prosocial phenotype", has been replicated across a wide variety of samples and cultures (Chierchia, Lesemann, Snower, Vogel, & Singer, 2017; Peysakhovich et al., 2014; Reigstad, Strømland, & Tinghög, 2017; Yamagishi et al., 2013).

The correlation between political views and the prosocial phenotype provides further evidence for a dispositional basis for political ideology. However, this correlation raises the question of whether and how these variables causally influence one another. At least three causal models are compatible with a cross-sectional correlation between political views and

the prosocial phenotype.

Prior work has tended to adopt a *prosociality-as-outcome* model. Under this model, biological and environmental factors interact to produce political views (Duckitt & Sibley, 2009) and people's political views then influence how they behave in economic games. This causal model is often assumed *a priori* to explain cross-sectional correlations between politics and behaviour in economic games. For example, Grünhage and Reuter (2020) write that "political orientation *predisposes for* a more trusting or cooperative behavior" (italics added; p. 22).

Alternatively, under the *prosociality-as-antecedent* model, biological and environmental factors interact to produce the behavioural predispositions captured by economic games, and these behavioural predispositions influence the expression of particular political views. This causal model is predicted by the dual evolutionary framework of political ideology (Claessens, Fischer, Chaudhuri, Sibley, & Atkinson, 2020), which explains political attitudes as shaped in part by basic social drives that were favoured during human evolution. Human group living evolved via two key shifts (Tomasello, Melis, Tennie, Wyman, & Herrmann, 2012); a shift towards increased cooperation with others, and a shift towards conformity to and enforcement of group-wide social norms. According to the dual evolutionary framework, variation in general drives for cooperation and group conformity arises from the interaction between heritable individual differences and socio-ecological environments. These general drives for cooperation and group conformity, together with individuals' immediate social context, produce variation in two dimensions of political ideology, often referred to as economic and social ideology (Claessens et al., 2020). This causal pathway — from behavioural predispositions to politics — is captured by Van Lange, Bekkers, Chirumbolo, and Leone (2012), who write that "political preferences and voting are *partially rooted in* interpersonal orientations" (italics added; p. 469).

Finally, the *common-cause* model predicts that both behaviour and political views

are caused by the same biological and environmental factors, but do not directly influence one another over time. This model is inspired by recent longitudinal evidence showing that personality does not causally precede political ideology, but instead personality and political ideology develop together in parallel, likely due to common causes from biological and environmental factors (Osborne & Sibley, 2020). Similarly, the common-cause model predicts that political ideology and the prosocial phenotype will be correlated, not because they influence one another over time, but because they share the same biological and environmental causes. This model is also consistent with the dual evolutionary framework of political ideology, insofar as heritable individual differences and socio-ecological environments influence both prosocial behaviour and political ideology simultaneously.

Figure 1 provides an overview of these causal models. All three models predict a cross-sectional correlation between political ideology and prosocial behaviour. As such, previous cross-sectional work cannot distinguish between them. Previous longitudinal work has shown that prosocial dispositions predict voting outcomes four weeks and eight months later (Van Lange et al., 2012) but, without concurrent measures of political views, this result is unable to directly support the prosociality-as-antecedent model.

Longitudinal panel data allow us to generate specific hypotheses. If the prosociality-as-outcome model is correct, then political views at time t should predict the prosocial phenotype at time $t + 1$, but the prosocial phenotype at time t should be unrelated to political views at time $t + 1$. If the prosociality-as-antecedent model is correct, the opposite should be true: the prosocial phenotype at time t should predict political views at time $t + 1$, but political views at time t should be unrelated to the prosocial phenotype at time $t + 1$. We further test whether a common cause is generating the observed relationships by controlling for a wide range of plausible demographic and personality confounds.

Here, we test between different causal hypotheses using a pre-registered cross-lagged
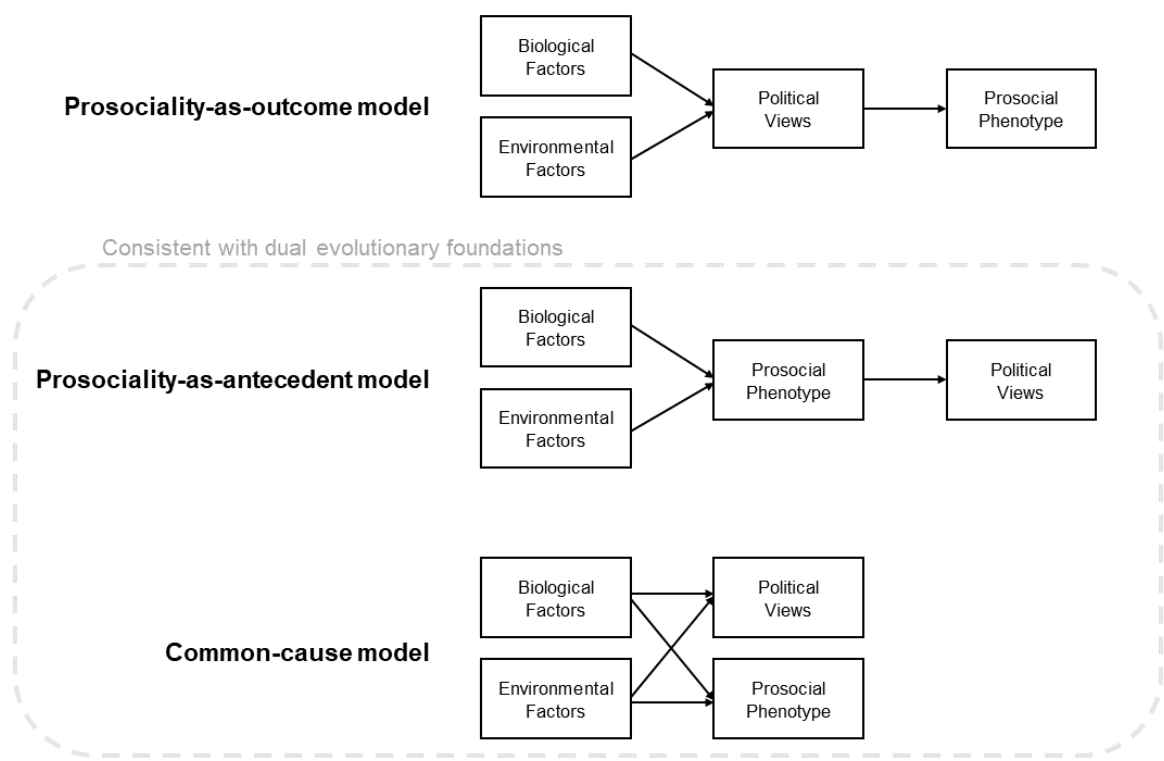
Figure 1. At least three theoretical causal models are compatible with the cross-sectional correlation between political views and the prosocial phenotype.

longitudinal panel design with two time-points separated by eighteen months. We estimate the directions of causality between the prosocial phenotype and several political views, including Social Dominance Orientation, views on economic issues, and political party support.

Methods

Ethical approval

Ethical approval was granted by REDACTED (ref: 021666). The study was performed in accordance with all the relevant guidelines and regulations. Participation was voluntary and informed consent was obtained from all participants prior to the study.

Participants and sampling

Participants were sampled from the New Zealand Attitudes and Values Study (NZAVS), an annual longitudinal self-report study that has been active since 2009. 1045 participants took part in the first wave of economic game data collection, were successfully paid for the study, and did not time out of their session. In the second wave, this sample size dropped to 631 (60% retention rate). We only analysed data from participants who completed both waves (n = 631; 411 females; mean age = 51 years; age range = 24 - 71 years). This sample size was the largest available to us and was not determined through a formal power analysis. These data are the same as used in our previous work (Claessens, Kelly, Sibley, Chaudhuri, & Atkinson, 2022; Claessens et al., 2023).

Materials

Self-report measures. Main self-report measures were taken from Waves 10 and 11 of the NZAVS. The primary measures of interest for this study were: Social Dominance Orientation (Pratto et al., 1994); support for income redistribution ("redistributing money and wealth more evenly among a larger percentage of the people in New Zealand through heavy taxes on the rich"); income attribution ("if incomes were more equal, people would be less motivated to work hard"); and support for New Zealand's centre-right National Party. We chose these measures because they all exhibited cross-sectional correlations with economic game behaviour in previous research (Claessens et al., 2023).

Other time-invariant covariates that could plausibly act as confounding variables were taken from Wave 10 of the NZAVS (Supplementary Table 1). These included socio-demographic variables (age, gender, ethnicity, education, measures of local deprivation, socio-economic status, and religiosity), personality variables (extraversion, agreeableness, conscientiousness, neuroticism, openness-to-experience, honesty-humility, and narcissism) (Sibley et al., 2011) and Right Wing Authoritarianism (Altemeyer, 1981).

We do not include RWA as a primary measure of interest in this study because our theoretical framework does not predict a longitudinal (or cross-sectional) relationship between authoritarian ideology and the prosocial phenotype (Claessens et al., 2020). Consistent with this, a prior analysis found no cross-sectional relationship between these variables (Claessens et al., 2023).

Battery of economic games.   Participants completed several incentivised one-shot economic games, conducted online in real-time using oTree (Chen, Schonger, & Wickens, 2016). These games measure prosocial behaviour and are largely identical to the games used in Peysakhovich et al. (2014). We used the strategy method to elicit responses in all roles. Participants played for points, which were converted to New Zealand dollars (1 point = $0.035).

The games measuring prosociality (paying a cost to benefit others) in our pre-registered analyses are as follows:

- *Dictator Game.* Player A is given 100 points. They must decide how many of these points to transfer to Player B. Player A keeps the remaining points. Player B is passive in the interaction.
- *Trust Game.* Players A and B both start with 50 points. First, Player A decides whether or not to transfer all 50 points to Player B, in the knowledge that the transferred amount will be tripled to 150 points. If Player A transfers, Player B now has 200 points. Player B must then decide to transfer 0 - 150 points back to Player A. There are thus two decisions in this game: giving and returning.
- *Public Goods Game.* Four players begin with 100 points each. They can contribute 0 - 100 points into a shared group project. All four decisions are made simultaneously, and then the amount in the group project is doubled and distributed evenly between all four players. Each player ends the game with their share from the group project, plus the points they initially refrained from contributing.

In further analyses that were not pre-registered, we tested the generalisability of our results by also including games where prosocial behaviour was elicited under the threat of future punishment (the Ultimatum Game and Second-Party Punishment Game). See Supplementary Methods for game descriptions and further details.

Procedure

The NZAVS collects self-report data both online and via paper surveys. In Wave 10 of the study, most of the data were collected between November 2018 and September 2019, and in Wave 11 of the study, most of the data were collected between October 2019 and November 2020. Data collection for the first wave of economic games was conducted between 18th February and 25th July 2019, and data collection for the second wave was conducted between 19th October and 11th November 2020 (Supplementary Figure 1).

Economic game sessions were conducted in real-time on midweek evenings, during which the games were completed in a randomised order. Participants were paid a $20 NZD show-up fee, plus their bonus payment. In the first wave, participants earned a bonus payment of $25.27 on average (SD = $2.52). In the second wave, participants earned a bonus payment of $21.39 on average (SD = $2.63).

Pre-registration

We pre-registered our hypotheses on the Open Science Framework (https://osf.io/ksw3x/?view_only=597d9d552bc142f4a6d59e4eba97b425) before running the second wave of economic game data collection. We deviated from our pre-registration in our interpretation of a lack of cross-lagged effects. In the pre-registration, we stated that a lack of cross-lagged effects, along with the presence of within-wave correlations, would support the common-cause model in Figure 1. However, an anonymous reviewer pointed out that this is not necessarily true, especially if potential common causes are not

statistically controlled for. We no longer make this statistical inference, but test several versions of the common cause model by systematically controlling for a range of potential confounds, including demographic and personality variables.

Statistical analysis

We used confirmatory factor analysis and structural equation modelling to test our pre-registered hypotheses. For measurement invariance analyses, we fitted a confirmatory factor analysis model with correlated item errors across waves to deal with non-independence of observations. For our longitudinal modelling, we used two-wave two-variable cross-lagged panel models.

86% of participants had complete data for all variables (see Supplementary Figure 2). To deal with missing data, we used multiple imputation with predictive mean matching (van Buuren & Groothuis-Oudshoorn, 2011). Multiple imputation procedures generate a number of plausible complete datasets with imputed missing values. Statistical models are then fitted to all imputed datasets and results are pooled across models to account for uncertainty in imputed values. We followed the advice of imputing at least as many datasets as the percentage of cases that are incomplete (Von Hippel, 2009). We therefore pooled our estimates across 20 imputed datasets. Visual inspection confirmed the plausibility of imputed values (see Supplementary Figure 3). Results for all pre-registered models were unchanged when using listwise deletion (see Supplementary Material).

Transparency and openness

The consent form for the NZAVS states that data will only be shared with other researchers on a case-by-case basis for the purposes of checking published work. For this reason, ethical concerns prevent us from making the dataset from this study publicly available. However, data are available on request from the lead author.

Pre-registration, analysis plan, R code for analyses, and Python code for the economic games are available at https://osf.io/ksw3x/?view_only=597d9d552bc142f4a6d59e4eba97b425. All analyses were conducted in R version 4.2.1 (R Core Team, 2019) using the *lavaan* package (Rosseel, 2012). Figures were created using *ggraph* (Pedersen, 2020), *cowplot* (Wilke, 2019), and *ggplot2* (Wickham, 2016) packages, multiple imputation was implemented using the *mice* package (van Buuren & Groothuis-Oudshoorn, 2011), reproducibility of all analyses was ensured by using the *targets* package (Landau, 2021), and the manuscript was generated using the *papaja* package (Aust & Barth, 2020).

Results

In the first step of our pre-registered analyses, we focused on the second wave of data in order to replicate our previous findings from the first wave (Claessens et al., 2023). First, we fitted a confirmatory factor analysis model with the Trust Game (Give), Trust Game (Return), Dictator Game, and Public Goods Game loading onto a "prosocial phenotype" latent variable. Supporting our pre-registered hypothesis, all factor loadings were significantly positive ($p < 0.05$) and the model fitted the data well (CFI = 0.99, RMSEA = 0.07, SRMR = 0.04; Supplementary Figure 4) (Hu & Bentler, 1999; MacCallum, Browne, & Sugawara, 1996). The games showed acceptable reliability in both the first wave ($\alpha = 0.61$, $\omega = 0.66$) and the second wave ($\alpha = 0.64$, $\omega = 0.70$), and the zero-order correlations between games were all positive and small-to-medium in size ($r = 0.20 - 0.37$; Supplementary Figure 5) supporting the existence of a single latent construct. These results held when we included prosocial behaviour in the Ultimatum Game and the Second-Party Punishment Game (see Supplementary Results).

Second, we fitted a structural equation model with SDO as the sole predictor of the prosocial phenotype latent variable. Supporting our pre-registered hypothesis, we found that SDO significantly negatively predicted the prosocial phenotype (unstandardised $b =$

-0.13, 95% confidence interval [-0.20 -0.06], $p < .001$; Supplementary Figure 6). This result held when we included prosocial behaviour in the Ultimatum Game and the Second-Party Punishment Game (see Supplementary Results). Together, these findings replicate our previous work with the same sample of participants eighteen months later (Claessens et al., 2023).

In the next step of our pre-registered analyses, we tested the measurement invariance of the prosocial phenotype latent variable across the two waves. Longitudinal measurement invariance testing ensures that latent factor structures are stable over time, an important prerequisite to cross-lagged panel modelling. We found support for strict invariance of the prosocial phenotype latent variable over time (see Supplementary Results) suggesting that the prosocial phenotype was measured with the same reliability over an eighteen-month period. Moreover, the between-wave latent correlation from the strict invariance model ($r = 0.65$, 95% CI [0.56 0.74]) suggested that the prosocial phenotype exhibited good test-retest reliability over time. These results held when we included prosocial behaviour in the Ultimatum Game and the Second-Party Punishment Game (see Supplementary Results).

We then proceeded to fit our pre-registered two-variable two-wave cross-lagged panel models. We first fitted our primary cross-lagged panel model, modelling the relationship between SDO and the prosocial phenotype over time (Figure 2a). We found significantly positive autoregressive effects: SDO in the first wave predicted SDO in the second wave (standardised $\beta = 0.81$, unstandardised $b = 0.84$, 95% CI [0.79 0.89], $p < .001$) and the prosocial phenotype in the first wave predicted the prosocial phenotype in the second wave ($\beta = 0.65$, $b = 0.73$, 95% CI [0.60 0.86], $p < .001$). Additionally, we found that the prosocial phenotype in the first wave negatively predicted SDO in the second wave ($\beta = -0.09$, $b = -0.14$, 95% CI [-0.24 -0.04], $p = .008$), but SDO in the first wave did not predict the prosocial phenotype in the second wave ($\beta = 0.00$, $b = 0.00$, 95% CI [-0.03 0.03], $p = .967$). These cross-lagged paths were significantly different from one another (difference in

unstandardised estimates $= 0.14$, 95% CI [0.03 0.24], $p = .011$). These results were unchanged when using listwise deletion rather than multiple imputation (Supplementary Figure 7) and when including the Ultimatum Game and Second-Party Punishment Game (see Supplementary Results).

According to Adachi and Willoughby (2015), even small cross-lagged effects in autoregressive models can be meaningful if the bivariate between-wave correlation is sufficiently large and the outcome variable is highly stable over time. Indeed, we find that the bivariate correlation between the prosocial phenotype in the first wave and SDO in the second wave is medium in size ($r = -0.29$) and SDO is highly stable over time ($\beta = 0.65$), suggesting that the cross-lagged effect of prosociality on future SDO is meaningful despite its smaller size.

To test possible common-cause explanations for this pattern of results, we ran additional cross-lagged panel models controlling for a range of time-invariant covariates that could plausibly act as confounds (Figures 2b and 2c; Supplementary Figures 7b and 7c). The cross-lagged path from the prosocial phenotype to later SDO remained significantly negative when controlling for most demographic and personality variables, but was attenuated when controlling for gender and ethnicity. Given these results, we ran exploratory multi-group cross-lagged panel models with separate groups for (1) male and female participants, and (2) participants of European ancestry and participants not of European ancestry (due to small sample sizes in individual Asian, Māori, and Pacific groups). These follow-up models revealed that the cross-lagged path from the prosocial phenotype to later SDO was significantly negative for males ($\beta = -0.12$, $b = -0.18$, 95% CI [-0.35 0.00], $p = .044$) but not for females ($\beta = -0.07$, $b = -0.10$, 95% CI [-0.22 0.02], $p = .098$) though these were both in the same direction. Similarly, the cross-lagged path from the prosocial phenotype to later SDO was significantly negative for participants of European ancestry ($\beta = -0.08$, $b = -0.13$, 95% CI [-0.25 -0.01], $p = .028$) but not for other participants ($\beta = -0.13$, $b = -0.17$, 95% CI [-0.37 0.02], $p = .079$) though these were both in
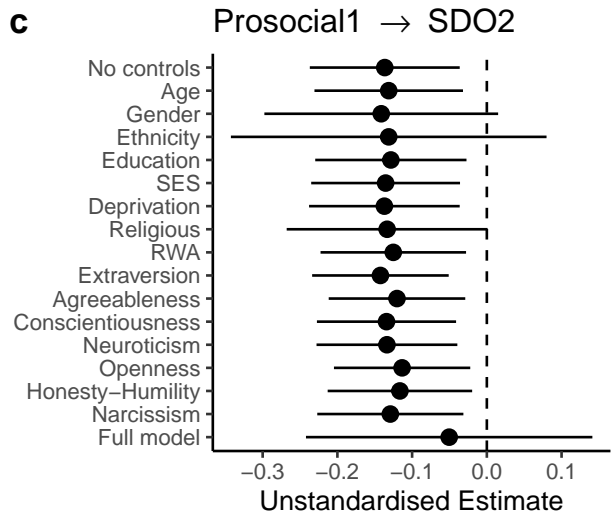
the same direction.

a
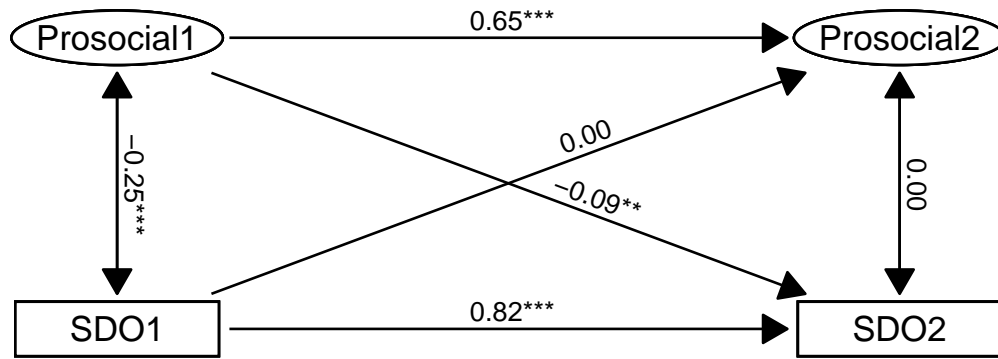
b
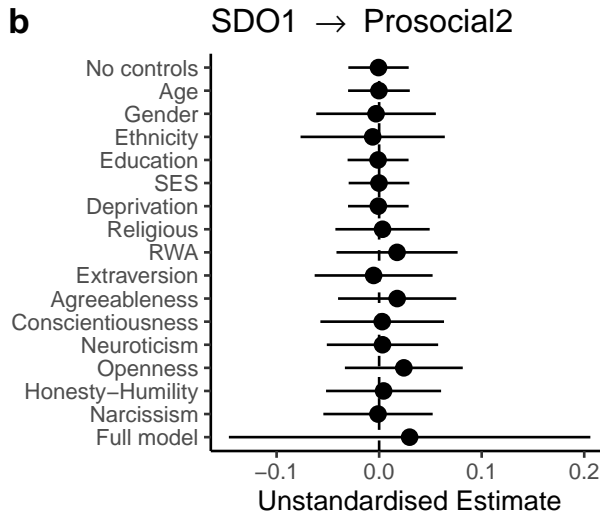
c

Figure 2. *Results of cross-lagged panel model with the prosocial phenotype and SDO, pooling over 20 imputed datasets.* (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, $***p < 0.001$, $**p < 0.01$, $*p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.

We also collected data on other self-reported measures of political views, including views on income redistribution, income attribution, and support for the centre-right

National Party. To assess the generalisability of our results, we swapped out SDO in our cross-lagged panel model for these additional measures.

When we included support for income redistribution in the model instead of SDO, we found a similar pattern of results: the prosocial phenotype positively predicted future support for income redistribution, but support for income redistribution did not predict future prosociality (Supplementary Figures 8 and 9). However, this pattern of results was attenuated when controlling for gender, ethnicity, religiosity, agreeableness, openness-to-experience, and honesty-humility, and when including the Ultimatum Game and the Second-Party Punishment Game in the measurement model for the prosocial phenotype (see Supplementary Results).

When we included income attribution views and support for New Zealand's centre-right National Party in the model instead of SDO, we initially found no cross-lagged effects over time (Supplementary Figures 10 – 13). However, we found that the prosocial phenotype positively predicted future support for the National Party when including the Ultimatum Game and Second-Party Punishment Game in the measurement model for the prosocial phenotype (see Supplementary Results).

Discussion

In a cross-lagged longitudinal analysis of prosocial behaviour and self-reported political attitudes, we have shown that the prosocial phenotype predicts SDO a year and a half later, but SDO does not predict future variation in the prosocial phenotype. This result is in line with the prosociality-as-antecedent model (Figure 1) which posits that general behavioural predispositions like the prosocial phenotype causally influence political views, but not vice versa. This causal model explains previously reported negative cross-sectional correlations between prosociality and SDO (Claessens et al., 2023; Haesevoets et al., 2015; Halali et al., 2018; Thielmann et al., 2020) as arising from a causal

relationship from behavioural preferences to political views.

Additionally, the cross-lagged path from prosociality to future SDO was robust to a wide range of time-invariant demographic and personality covariates. It is therefore unlikely that these variables are acting as common causes and generating the observed pattern of relationships (though, of course, it remains possible that other unmodelled common causes could exist). However, the cross-lagged path from prosociality to future SDO was attenuated when controlling for gender and ethnicity. In particular, exploratory analyses revealed that the cross-lagged effect held only for male participants of European ancestry. This is similar to the finding that upper body strength is related to support for inequality in males, but not females (Petersen & Laustsen, 2019). A possible evolutionary explanation for this effect of gender might be that humans have large sexual dimorphism in strength and formidability and, therefore, a correlation between the prosocial phenotype and dominative/competitive tactics for resource distribution is more likely among males. But this does not explain the effect of ethnicity. Perhaps a simpler explanation is that SDO is generally higher among males and people from dominant ethnic groups (Pratto et al., 1994; Sidanius, Levin, Liu, & Pratto, 2000). Among these participants, there is potentially more room for a change in SDO over time, whereas female participants from minority ethnic groups may have already hit the floor of the scale and therefore have less room for change. In line with this explanation, when we look at the differences in SDO between the two waves, we find that these difference scores have a higher variance for males of European ancestry (variance = 0.39) compared to other participants (variance = 0.24; Levene's test, $F(1,564) = 13.54$, $p < .001$) suggesting that SDO had more room for change over time among males of European ancestry. Future research should disentangle the roles of gender and ethnicity in the causal relationship between behavioural predispositions and political views.

When we generalised our cross-lagged analysis to other political views, we found more mixed results. As expected, the prosocial phenotype positively predicted future support for

income redistribution, though this effect was attenuated in some models. However, in our initial models, the prosocial phenotype was not longitudinally related to income attribution beliefs or support for New Zealand's centre-right National party. These null results could arise from a shared cause influencing these variables at the same time (i.e., the common-cause model in Figure 1) creating a cross-sectional relationship but no longitudinal effects. However, this pattern of null results is also consistent with other explanations. For example, it may be that party affiliation and income attribution beliefs are simply less amenable to change over time: people rarely shift their political party affiliation (Pew Research Center, 2020) and income attribution beliefs have been characterised as a stable individual difference (Osborne & Weiner, 2015; but see Piff et al., 2020). Alternatively, cross-lagged effects may exist between the prosocial phenotype and these variables but on longer or shorter timescales than studied here. It is therefore difficult to interpret these null results as supporting any particular causal model from Figure 1. Moreover, although we found that the prosocial phenotype positively predicted future support for the National Party when including the full set of games in the measurement model (Supplementary Results), this finding is also difficult to interpret given the negative bivariate correlation between these variables in the data ($r = $ -0.18).

Nevertheless, our results for SDO and views on income redistribution are broadly consistent with the dual evolutionary foundations framework for political ideology. The cross-lagged effects of the prosocial phenotype on these variables support either the prosociality-as-antecedent or common cause model, and none of our results directly support the causal model that is inconsistent with the dual evolutionary foundations (the prosociality-as-outcome model). Moreover, our measurement invariance analyses and the latent correlation from our strict invariance model revealed that the prosocial phenotype latent variable had adequate test-retest reliability over eighteen months, supporting another central claim from the dual evolutionary foundations framework that the general social drives that partly shape political ideology should be relatively stable over time. This

finding expands on studies showing that prosocial behaviour in individual economic games positively covaries when measured over four months (Peysakhovich et al., 2014; Reigstad et al., 2017), one year (Lönnqvist, Verkasalo, Walkowitz, & Wichardt, 2015), and even six years (Carlsson, Johansson-Stenman, & Nam, 2014).

One important limitation of this study is our use of two-wave cross-lagged panel models to determine longitudinal effects. Cross-lagged panel models have been criticised for not correctly partitioning within-person change from stable between-person differences (Hamaker, Kuiper, & Grasman, 2015). As an alternative to the cross-lagged panel model, Hamaker et al. (2015) proposed the random-intercept cross-lagged panel model, which estimates a random intercept to capture stable individual differences over time. Unfortunately, random-intercept cross-lagged panel models require at least three waves of data to be identified (Hamaker et al., 2015) and so we were limited by our two waves of data in this study. Moving forward, it will be vital to extend this study to include additional waves of behavioural and self-report data to determine if our results hold when accounting for stable between-person differences.

Future research should expand our multi-wave behavioural and self-report approach to include additional measures. In particular, to provide a complete picture for the dual evolutionary foundations framework, future research should test whether a "group conformist phenotype" predicts future variation in RWA and social policy views (Claessens et al., 2020). We did not focus on RWA in the current study as RWA is not theorised to be related to the prosocial phenotype. Instead, our previous work suggests that RWA covaries cross-sectionally with conformist behaviour in the rule following task and social learning tasks (Claessens et al., 2023; Fischer, Atkinson, & Chaudhuri, 2021). Since these studies only measured conformist behaviour in single time slices, we were unable to extend this work longitudinally to make causal claims. Such longitudinal studies have the potential to unpack the causal pathways underlying the full diversity of human political ideology.

Acknowledgements

Data Availability

The data described in the paper are part of the New Zealand Attitudes and Values Study (NZAVS). Full copies of the NZAVS data files are held by all members of the NZAVS management team and advisory board. A de-identified dataset containing the variables analysed in this manuscript is available upon request from the corresponding author, or any member of the NZAVS advisory board, for replication purposes.

Competing Interests

The authors declare no competing interests.

References

Adachi, P., & Willoughby, T. (2015). Interpreting effect sizes when controlling for stability effects in longitudinal autoregressive models: Implications for psychological science. *European Journal of Developmental Psychology*, *12*(1), 116–128. https://doi.org/10.1080/17405629.2014.963549

Altemeyer, B. (1981). *Right-wing authoritarianism.* University of Manitoba Press.

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown.* Retrieved from https://github.com/crsh/papaja

Brandt, M. J., Reyna, C., Chambers, J. R., Crawford, J. T., & Wetherell, G. (2014). The ideological-conflict hypothesis: Intolerance among both liberals and conservatives. *Current Directions in Psychological Science*, *23*(1), 27–34. https://doi.org/10.1177/0963721413510932

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter.* Chicago: Wiley.

Carlsson, F., Johansson-Stenman, O., & Nam, P. K. (2014). Social preferences are stable over long periods of time. *Journal of Public Economics*, *117*, 104–114. https://doi.org/https://doi.org/10.1016/j.jpubeco.2014.05.009

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree-an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Chierchia, G., Lesemann, F. P., Snower, D., Vogel, M., & Singer, T. (2017). Caring cooperators and powerful punishers: Differential effects of induced care and power motivation on different types of economic decision making. *Scientific Reports*, *7*(1), 11068.

Chirumbolo, A., Leone, L., & Desimoni, M. (2016). The interpersonal roots of politics: Social value orientation, socio-political attitudes and prejudice. *Personality and Individual Differences*, *91*, 144–153.

Claessens, S., Fischer, K., Chaudhuri, A., Sibley, C. G., & Atkinson, Q. D. (2020). The

    dual evolutionary foundations of political ideology. *Nature Human Behaviour*, *4*,

    336–345. https://doi.org/10.1038/s41562-020-0850-9

Claessens, S., Kelly, D., Sibley, C. G., Chaudhuri, A., & Atkinson, Q. D. (2022).

    Cooperative phenotype predicts climate change belief and pro-environmental behaviour.

    Scientific Reports, *12*(1), 12730.

Claessens, S., Sibley, C. G., Chaudhuri, A., & Atkinson, Q. D. (2023). Cooperative and

    conformist behavioural preferences predict the dual dimensions of political ideology.

    Scientific Reports, *13*(1), 4886.

Duckitt, J., & Sibley, C. G. (2009). A dual-process motivational model of ideology, politics,

    and prejudice. *Psychological Inquiry*, *20*(2-3), 98–109.

    https://doi.org/10.1080/10478400903028540

Fischer, K., Atkinson, Q. D., & Chaudhuri, A. (2021). Experiments in political psychology.

    In A. Chaudhuri (Ed.), *A research agenda for experimental economics*. Cheltenham:

    Edward Elgar.

Grünhage, T., & Reuter, M. (2020). Political orientation is associated with behavior in

    public-goods- and trust-games. *Political Behavior*, (0123456789).

    https://doi.org/10.1007/s11109-020-09606-5

Haesevoets, T., Reinders Folmer, C., Bostyn, D. H., & Van Hiel, A. (2018). Behavioural

    consistency within the prisoner's dilemma game: The role of personality and situation.

    European Journal of Personality, *32*(4), 405–426. https://doi.org/10.1002/per.2158

Haesevoets, T., Reinders Folmer, C., & Van Hiel, A. (2015). Cooperation in mixed-motive

    games: The role of individual differences in selfish and social orientation. *European

    Journal of Personality*, *29*(4), 445–458. https://doi.org/10.1002/per.1992

Halali, E., Dorfman, A., Jun, S., & Halevy, N. (2018). More for us or more for me? Social

    dominance as parochial egoism. *Social Psychological and Personality Science*, *9*(2),

    254–262. https://doi.org/10.1177/1948550617732819

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the
    cross-lagged panel model. *Psychological Methods*, *20*(1), 102.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure
    analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*,
    6(1), 1–55. https://doi.org/10.1080/10705519909540118

Inbar, Y., Pizarro, D., Iyer, R., & Haidt, J. (2012). Disgust sensitivity, political
    conservatism, and voting. *Social Psychological and Personality Science*, *3*(5), 537–544.

Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political ideology: Its structure,
    functions, and elective affinities. *Annual Review of Psychology*, *60*(1), 307–337.
    https://doi.org/10.1146/annurev.psych.60.110707.163600

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism
    as motivated social cognition. *Psychological Bulletin*, *129*(3), 339–375.

Landau, W. M. (2021). The targets R package: A dynamic Make-like function-oriented
    pipeline toolkit for reproducibility and high-performance computing. *Journal of Open
    Source Software*, *6*(57), 2959. Retrieved from https://doi.org/10.21105/joss.02959

Lönnqvist, J.-E., Verkasalo, M., Walkowitz, G., & Wichardt, P. C. (2015). Measuring
    individual risk attitudes in the lab: Task or ask? An empirical comparison. *Journal of
    Economic Behavior & Organization*, *119*, 254–266.
    https://doi.org/https://doi.org/10.1016/j.jebo.2015.08.003

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and
    determination of sample size for covariance structure modeling. *Psychological Methods*,
    1(2), 130–149. https://doi.org/10.1037/1082-989x.1.2.130

Osborne, D., & Sibley, C. G. (2020). Does openness to experience predict changes in
    conservatism? A nine-wave longitudinal investigation into the personality roots to
    ideology. *Journal of Research in Personality*, *87*, 103979.
    https://doi.org/10.1016/j.jrp.2020.103979

Osborne, D., & Weiner, B. (2015). A latent profile analysis of attributions for poverty:

Identifying response patterns underlying people's willingness to help the poor. *Personality and Individual Differences*, *85*, 149–154. https://doi.org/https://doi.org/10.1016/j.paid.2015.05.007

Pedersen, T. L. (2020). *ggraph: An implementation of grammar of graphics for graphs and networks.* Retrieved from https://CRAN.R-project.org/package=ggraph

Petersen, M. B., & Laustsen, L. (2019). Upper-body strength and political egalitarianism: Twelve conceptual replications. *Political Psychology*, *40*(2), 375–394. https://doi.org/https://doi.org/10.1111/pops.12505

Pew Research Center. (2020). *Voters rarely switch parties, but recent shifts further educational, racial divergence.* Retrieved from https://www.pewresearch.org/politics/2020/08/04/voters-rarely-switch-parties-but-recent-shifts-further-educational-racial-divergence/

Peysakhovich, A., Nowak, M. A., & Rand, D. G. (2014). Humans display a "cooperative phenotype" that is domain general and temporally stable. *Nature Communications*, *5*, 4939. https://doi.org/10.1038/ncomms5939

Piff, P. K., Wiwad, D., Robinson, A. R., Aknin, L. B., Mercier, B., & Shariff, A. (2020). Shifting attributions for poverty motivates opposition to inequality and enhances egalitarianism. *Nature Human Behaviour*, *4*(5), 496–505. https://doi.org/10.1038/s41562-020-0835-8

Pisor, A. C., Gervais, M. M., Purzycki, B. G., & Ross, C. T. (2020). Preferences and constraints: The value of economic games for studying human behaviour. *Royal Society Open Science*, *7*(6), 192090. https://doi.org/10.1098/rsos.192090

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*(4), 741.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from

https://www.R-project.org/

Reigstad, A. G., Strømland, E. A., & Tinghög, G. (2017). Extending the cooperative
    phenotype: Assessing the stability of cooperation across countries. *Frontiers in
    Psychology*, *8*, 1990. https://doi.org/10.3389/fpsyg.2017.01990

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of
    Statistical Software*, *48*(2), 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/

Sibley, C. G., Luyten, N., Purnomo, M., Mobberley, A., Wootton, L. W., Hammond, M.
    D., et al.others. (2011). The Mini-IPIP6: Validation and extension of a short measure
    of the Big-Six factors of personality in New Zealand. *New Zealand Journal of
    Psychology (Online)*, *40*(3), 142.

Sidanius, J., Levin, S., Liu, J., & Pratto, F. (2000). Social dominance orientation,
    anti-egalitarianism and the political psychology of gender: An extension and
    cross-cultural replication. *European Journal of Social Psychology*, *30*(1), 41–67.

Thielmann, I., Hilbig, B. E., & Zettler, I. (2022). The dispositional basis of human
    prosociality. *Current Opinion in Psychology*, *43*, 289–294.

Thielmann, I., Spadaro, G., & Balliet, D. (2020). Personality and prosocial behavior: A
    theoretical framework and meta-analysis. *Psychological Bulletin*, *146*(1), 30–90.
    https://doi.org/10.1037/bul0000217

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key
    steps in the evolution of human cooperation: The interdependence hypothesis. *Current
    Anthropology*, *53*(6), 673–692. https://doi.org/10.1086/668207

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by
    chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. Retrieved from
    https://www.jstatsoft.org/v45/i03/

Van Lange, P. A. M., Bekkers, R., Chirumbolo, A., & Leone, L. (2012). Are conservatives
    less likely to be prosocial than liberals? From games to ideology, political preferences
    and voting. *European Journal of Personality*, *26*(5), 461–473.

https://doi.org/10.1002/per.845

Von Hippel, P. T. (2009). How to impute interactions, squares, and other transformed variables. *Sociological Methodology, 39*(1), 265–291. https://doi.org/10.1111/j.1467-9531.2009.01215.x

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wilke, C. O. (2019). *cowplot: Streamlined Plot Theme and Plot Annotations for "ggplot2".* Retrieved from https://CRAN.R-project.org/package=cowplot

Yamagishi, T., Mifune, N., Li, Y., Shinada, M., Hashimoto, H., Horita, Y., et al.others. (2013). Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. *Organizational Behavior and Human Decision Processes, 120*(2), 260–271.

Zaller, J. R. (1992). *The nature and origins of mass opinion.* Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511818691

Supplementary Information

Prosocial phenotype predicts political views on hierarchy and redistribution eighteen months later

Scott Claessens[1], Chris G Sibley[1], Ananish Chaudhuri[2,3], & Quentin D Atkinson[1,4]

[1] School of Psychology, University of Auckland, New Zealand

[2] Department of Economics, University of Auckland, Auckland, New Zealand

[3] CESifo, Munich, Germany

[4] Max Planck Institute for the Science of Human History, Jena, Germany

Supplementary Methods

Descriptions for Ultimatum Game and Second-Party Punishment Game.   In
the main text, we test the robustness of our results to the inclusion of prosocial behaviour
in the following two games:

- *Ultimatum Game.* Player A starts with 100 points, and Player B starts with nothing.
  Player A must decide how many points to transfer to Player B. However, Player B
  simultaneously specifies their 'minimum acceptable offer': namely, the lowest transfer
  from Player A that they will accept. If Player A's transfer amount is lower than this
  minimum acceptable offer, both players end the game with 0 points. Otherwise,
  Player B receives the transfer amount, and Player A keeps the remaining points.

- *Second-Party Punishment Game.* Players A and B start with 100 points. This game
  has two stages: the transfer stage, and the penalty stage. In the transfer stage, each
  player decides whether to transfer 30 points to the other player. Any transferred
  points are doubled before the other player receives them. Decisions are made
  simultaneously. The transfer stage follows the payoff matrix of a Prisoner's Dilemma.
  Then, in the penalty phase, both players can pay 0–10 points to remove points from
  the other player, depending on their decision in the transfer stage. Each paid point
  removes 5 points from the other player.

Procedure for economic game sessions.   In both waves, participants were booked
into sessions on midweek evenings and completed the session online in real-time. Session
sizes varied between 14 and 130 participants. Although participants knew they were
completing the study with other participants from the New Zealand Attitudes and Values
Study, they did not know specifically who they were interacting with in the session or how
many other people there were in the session.

Participants completed a consent form before proceeding to the eight behavioural tasks. In
the first wave, all eight tasks were completed in a randomised order. In the second wave,

the economic games shared with the first wave were completed first in a randomised order, followed by two new tasks (the rule following and social information use tasks) which were presented in a separately randomised order. For each task, participants read the instructions for the task, completed a comprehension question, and then proceeded to make their decisions. Participants did not receive any information about their earnings over the course of the session to avoid any potential wealth or learning effects between games.

As well as the prosocial decisions listed in the main text, participants in both waves also made punishment decisions in several punishment games (Ultimatum Game, Third Party Punishment Game, and Second Party Punishment Game). Moreover, in the first wave, participants completed additional coordination games (Stag Hunt Game and Stag Hunt Game with Punishment) and, in the second wave, participants completed additional behavioural measures of rule following and social information use (Claessens et al., 2023).

After making their decisions for all the tasks, participants entered a waiting lobby in which they waited for all other participants in their session to complete the tasks. If participants took longer than 55 minutes to complete the tasks, they were skipped ahead to the waiting lobby. Timeouts were still paid their show-up fee, but not their bonus. In the first wave, participants took 22 minutes on average to complete all eight tasks (SD = 7 minutes, range = 9 - 47 minutes), and in the second wave, participants took 24 minutes on average (SD = 8 minutes, range = 9 - 49 minutes).

After the sessions, participants were randomly matched into groups to determine bonus payments. After being matched, participants were informed about their outcomes in each of the economic games.

Potential learning between waves.   To determine whether learning occurred between waves and affected subsequent game behaviour, we estimated differences in behaviour between the two waves (see Supplementary Table 2). For all games except the Dictator Game, we found no differences in prosocial behaviour between waves. In the Dictator

Game, participants gave slightly less in the second wave, but this difference was small (2 fewer points given, on average, out of 100). These results suggest that learning biases are not a major concern in our study.

Potential bias due to dropouts between waves. Comparisons of drop-outs and non-drop-outs suggested that retention did not systematically bias our sample (see Supplementary Table 3).

Supplementary Results

Pre-registered measurement invariance results. We tested for measurement invariance of the prosocial phenotype factor structure in a series of increasingly restrictive nested models. For all model comparisons, we pre-registered the use of changes in fit statistics as thresholds for diagnosing reduced model fit [ΔComparative Fit Index (CFI) < -0.01, ΔRoot Mean Square Error of Approximation (RMSEA) > 0.015] rather than χ^2 differences which are sensitive to large sample sizes. To deal with non-independence of observations, all measurement invariance models had correlated item errors across waves.

First, we fitted a configural invariance model, which freely estimated the two latent variables simultaneously (Supplementary Table 4). As expected, this configural invariance model fitted the data well (CFI = 0.99, RMSEA = 0.04) and all loadings were significantly positive. Second, we fitted a metric invariance model, which constrained the item loadings to equality across the two waves. Model fit did not substantially change (ΔCFI = -0.006, ΔRMSEA = 0.006). Third, we fitted a scalar invariance model, which constrained the item loadings, intercepts, and thresholds to equality across the two waves. Again, model fit did not substantially change (ΔCFI = 0.000, ΔRMSEA = -0.004). Fourth, and finally, we fitted a strict invariance model, which constrained the item loadings, intercepts, thresholds, and variances to equality across waves. Model fit remained unchanged (ΔCFI = 0.000, ΔRMSEA = -0.002). Measurement invariance analysis thus supports strict invariance of the prosocial phenotype latent variable over time.

Including the Ultimatum Game and Second-Party Punishment Game. In addition to our pre-registered confirmatory factor analysis model with the Trust Game (Give), Trust Game (Return), Dictator Game, and Public Goods Game, we additionally included the Ultimatum Game (Offer) and Second-Party Punishment Game (Give) to estimate the "prosocial phenotype" latent variable. As with our pre-registered analysis, all factor loadings were significantly positive ($p < 0.05$) and the model fitted the data well

(CFI = 0.96, RMSEA = 0.08, SRMR = 0.06; Supplementary Figure 14) (Hu & Bentler, 1999; MacCallum et al., 1996). The full set of games showed acceptable reliability in both the first wave ($\alpha = 0.66$, $\omega = 0.75$) and the second wave ($\alpha = 0.68$, $\omega = 0.74$), and the zero-order correlations between games were all positive and small-to-medium in size ($r = 0.12 - 0.39$; Supplementary Figure 15) supporting the existence of a single latent construct.

We fitted a structural equation model with SDO as the sole predictor of the prosocial phenotype latent variable, including the additional games. As with our pre-registered analysis, we found that SDO significantly negatively predicted the prosocial phenotype (unstandardised $b = $ -0.15, 95% confidence interval [-0.21 -0.08], $p < .001$; Supplementary Figure 16).

We also assessed measurement invariance for the prosocial phenotype when including the Ultimatum Game and Second-Party Punishment Game. As with our pre-registered analysis, we found support strict invariance of the prosocial phenotype over time (Supplementary Table 5).

Finally, we repeated our pre-registered cross-lagged panel models including the Ultimatum Game and Second-Party Punishment Game in the measurement model for the prosocial phenotype (Supplementary Figures 17 – 20). The results for SDO and income attribution beliefs were unchanged from the pre-registered models. However, in the model with additional games, the prosocial phenotype no longer predicted future support for income redistribution and *positively* predicted future support for the National Party. This latter result is difficult to square with the negative bivariate correlation between the prosocial phenotype in the first wave and support for the National party in the second wave ($r = $ -0.18).

Supplementary Figures
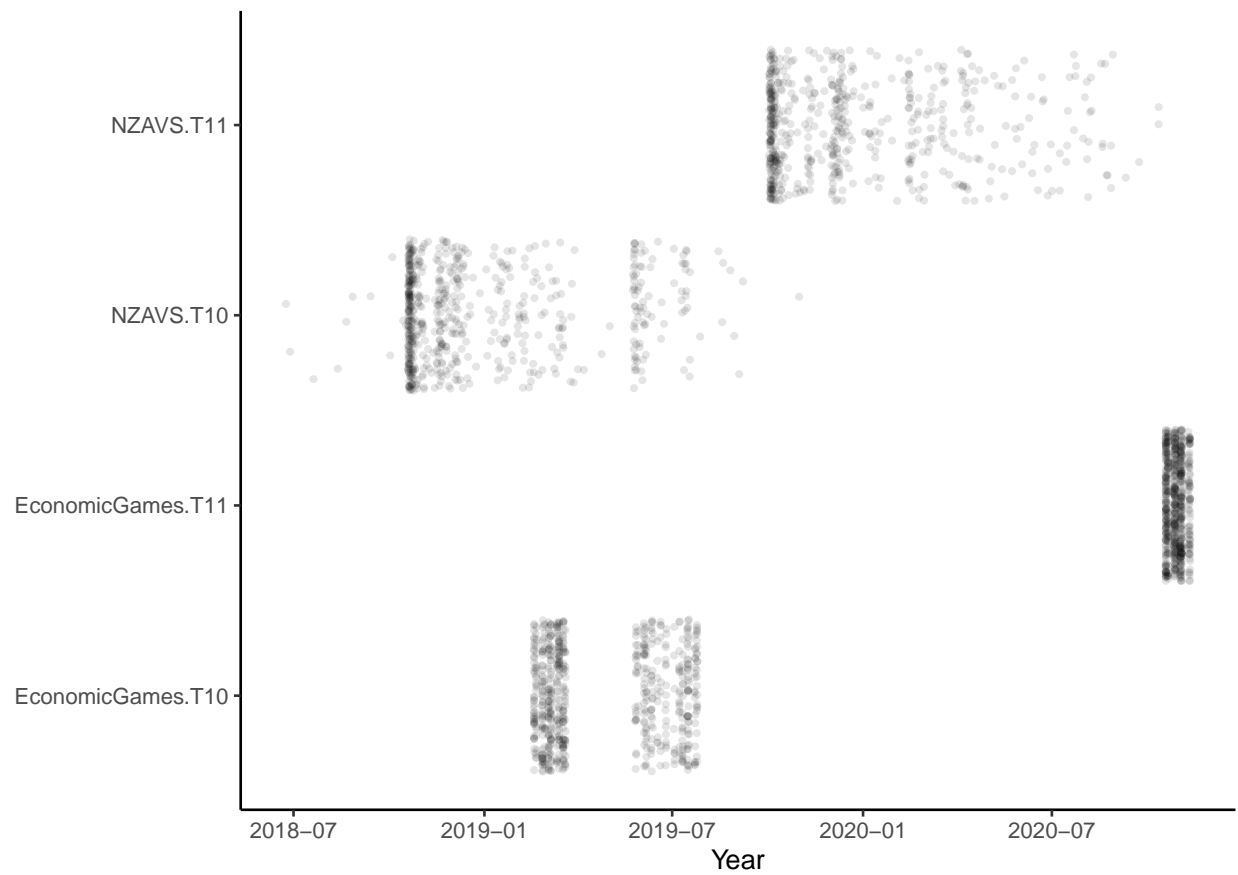
Supplementary Figure 1. Data collection timeline for NZAVS Wave 10, NZAVS Wave 11, and both waves of economic game data collection (n = 631). Each point is an individual participant. Note the break in data collection in February 2019 due to the Christchurch terrorist attack.
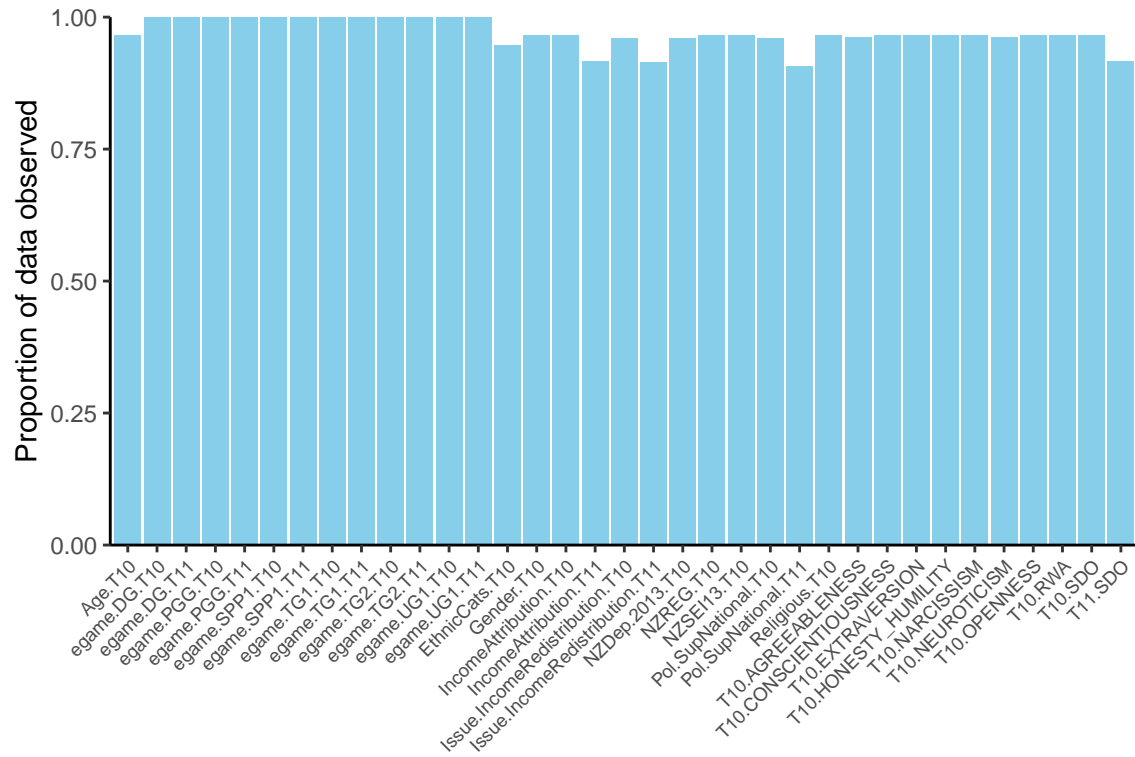
Supplementary Figure 2. Bar plot showing proportions of observed data for all variables included in the study. T10 = NZAVS Wave 10, T11 = NZAVS Wave 11.
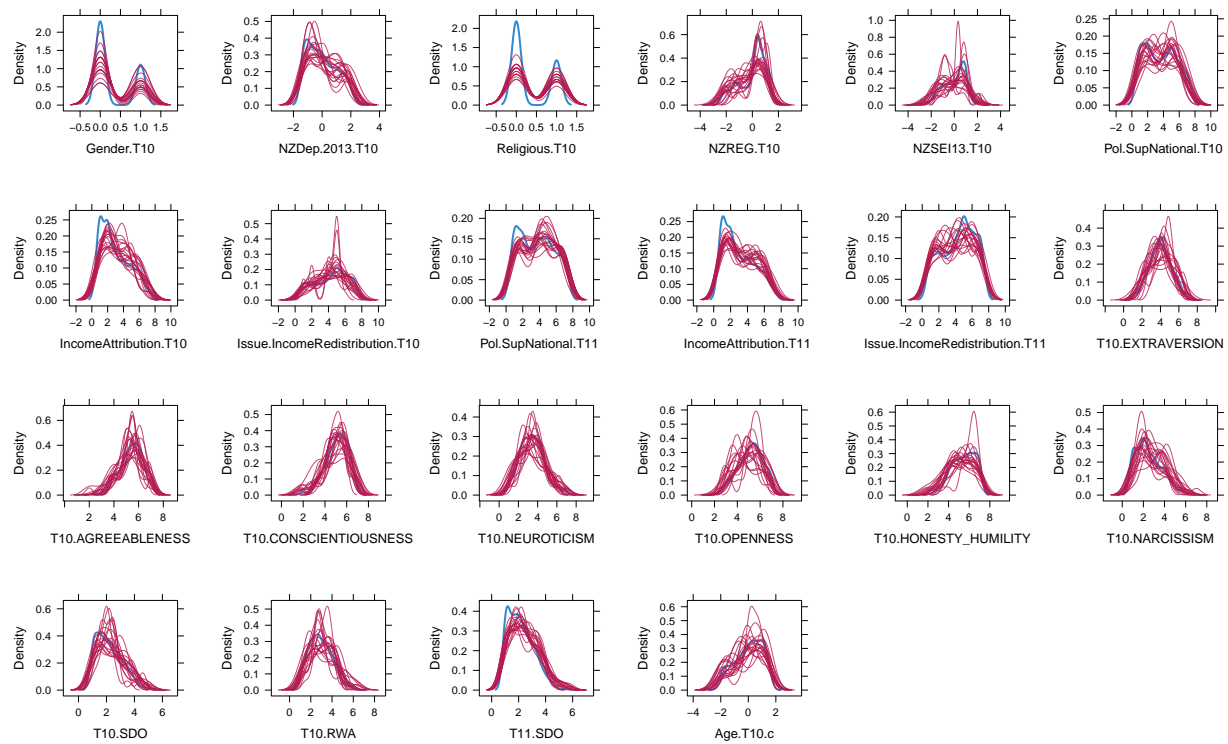
Supplementary Figure 3. Density plots showing imputed values from 20 multiply imputed datasets (pink) against observed values (blue). Data were imputed using predictive mean matching.
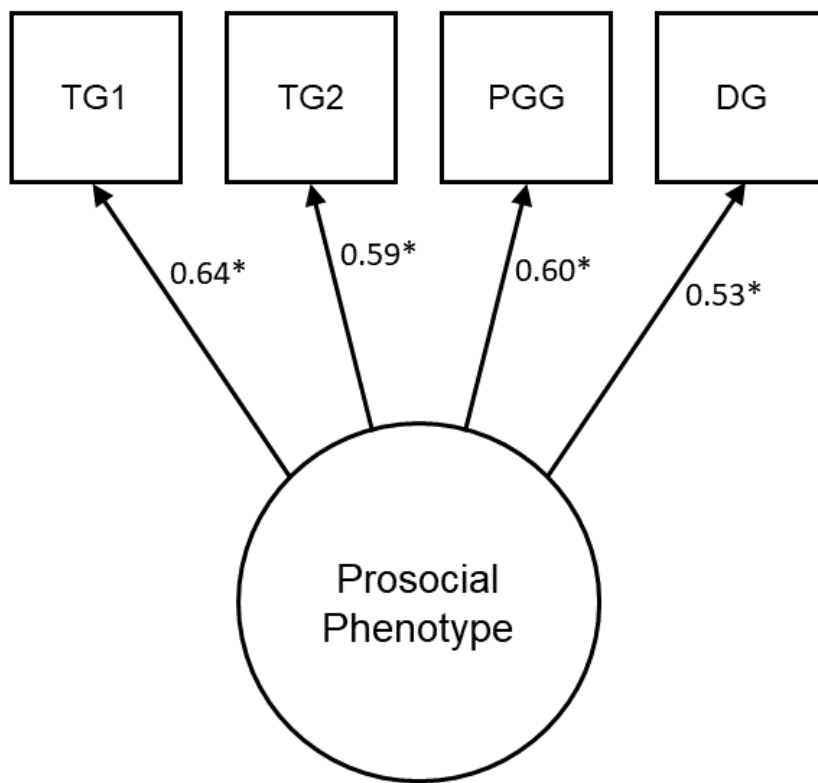
Supplementary Figure 4. *Confirmatory factor model for the prosocial phenotype in Wave 2.* TG1 is treated as a binary endogenous variable. Numbers are standardised coefficients. *p < 0.05. TG1 = Trust Game (Give), TG2 = Trust Game (Return), PGG = Public Goods Game, DG = Dictator Game.
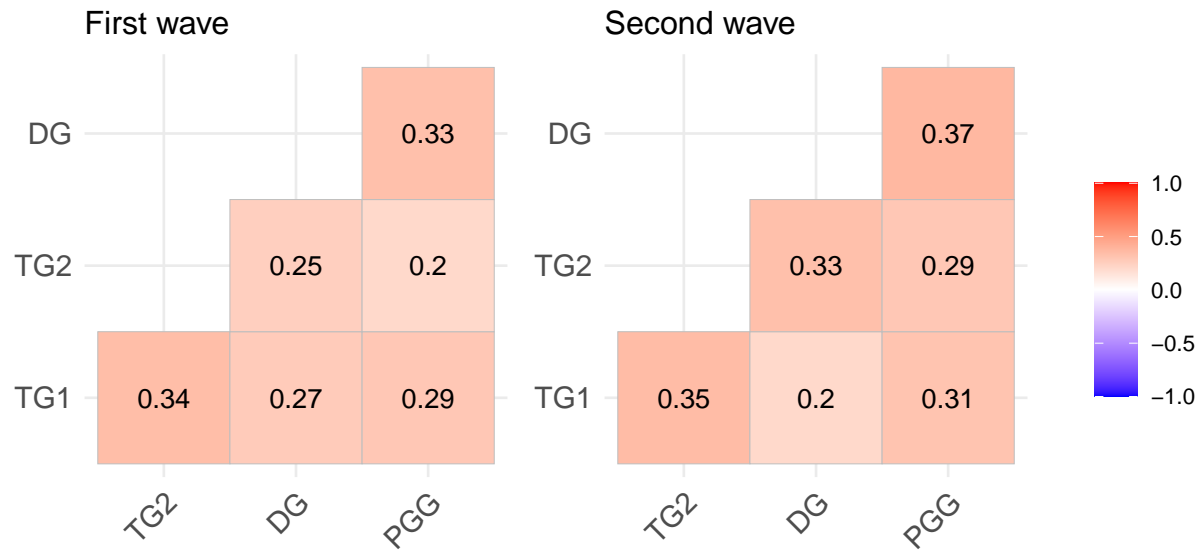
Supplementary Figure 5. *Zero-order correlations between game decisions in the first and second wave.* All p-values < 0.01. TG1 = Trust Game (Give), TG2 = Trust Game (Return), DG = Dictator Game, PGG = Public Goods Game.
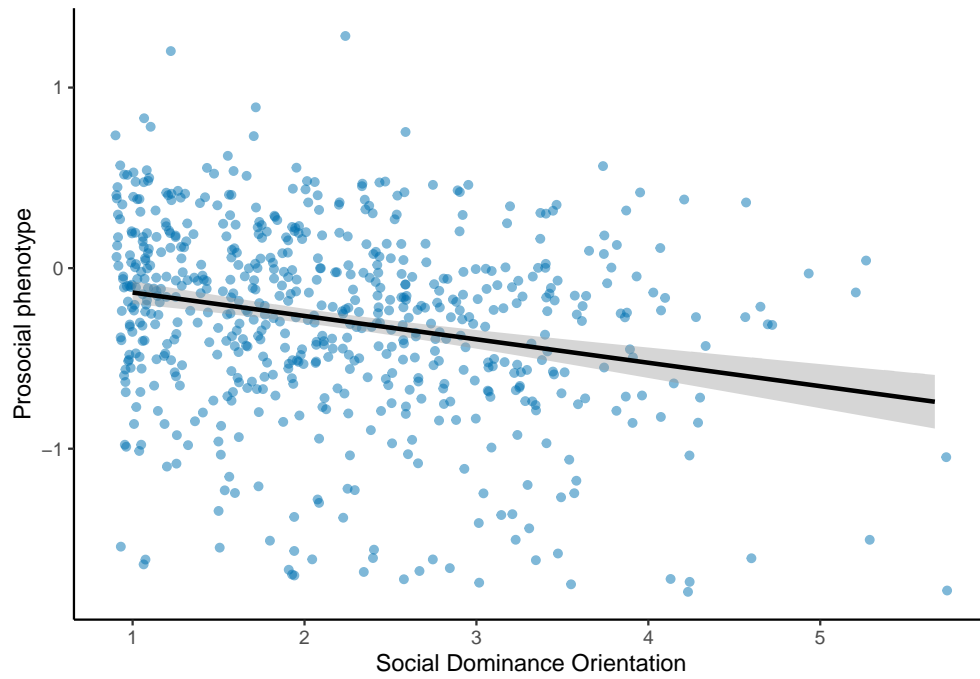
Supplementary Figure 6. Social Dominance Orientation (mean score) is negatively related to model-predicted prosocial latent variable scores in the second wave.
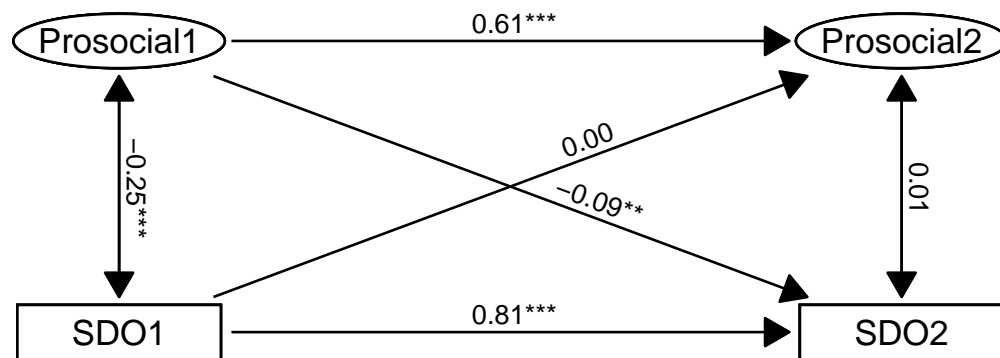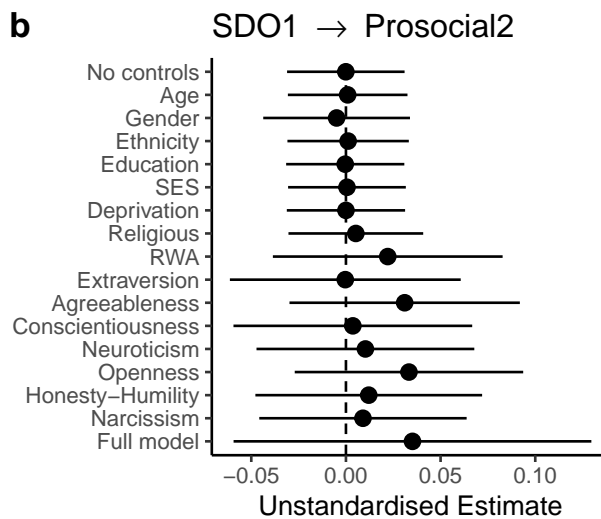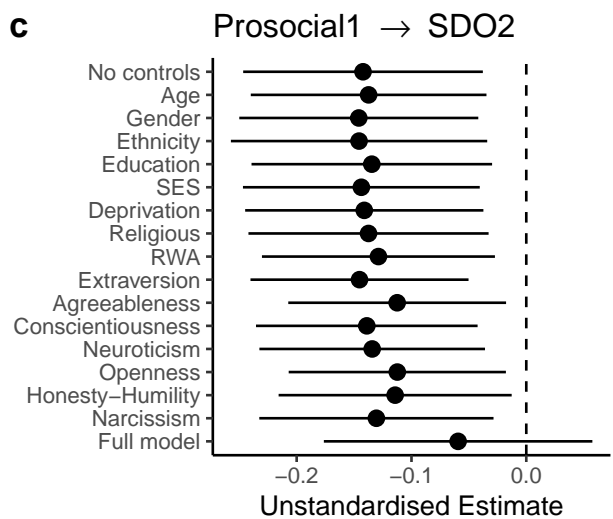
Supplementary Figure 7. Results of cross-lagged panel model with the prosocial phenotype and SDO, analysing listwise-deleted data. (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, **$p < 0.01$, *$p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.
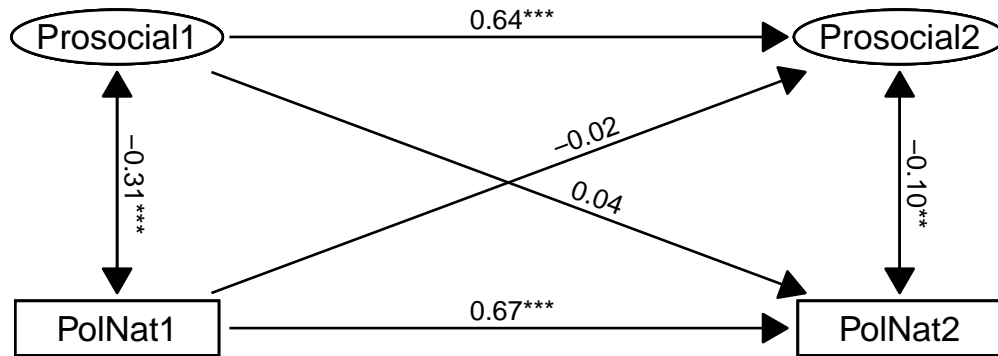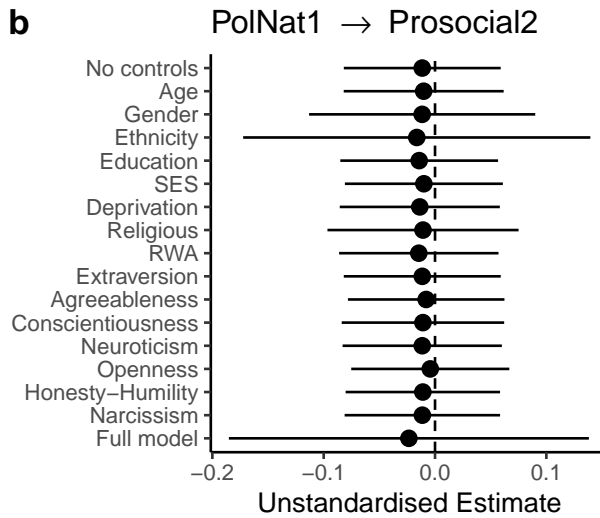
a

b IncRed1 → Prosocial2   **c** Prosocial1 → IncRed2
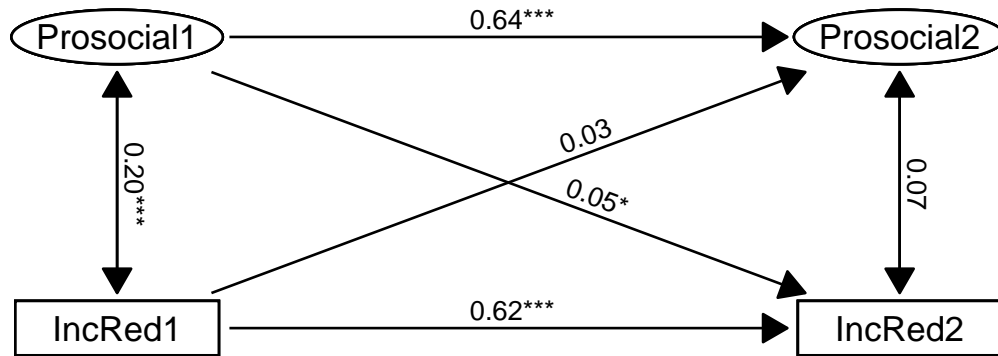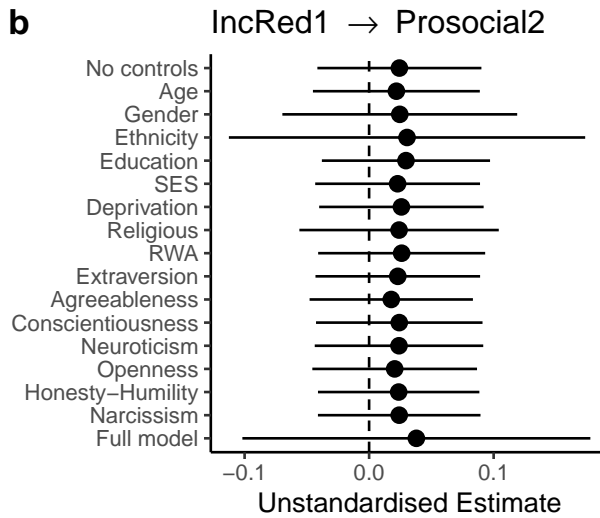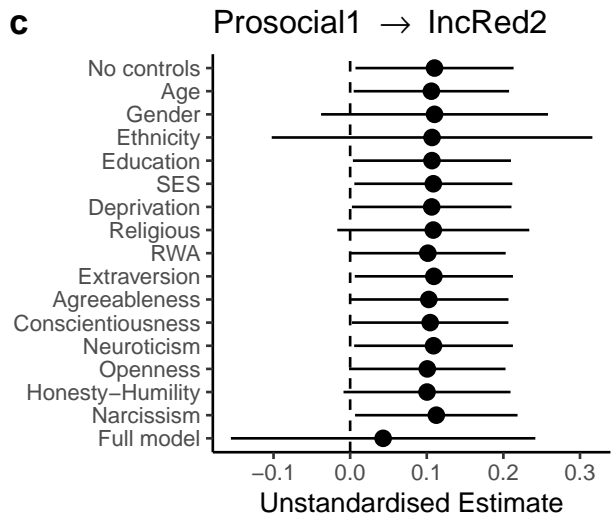
Supplementary Figure 8. Results of cross-lagged panel model with the prosocial phenotype and support for income redistribution, pooling over 20 imputed datasets. (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Support for income redistribution is treated as ordinal. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.

a

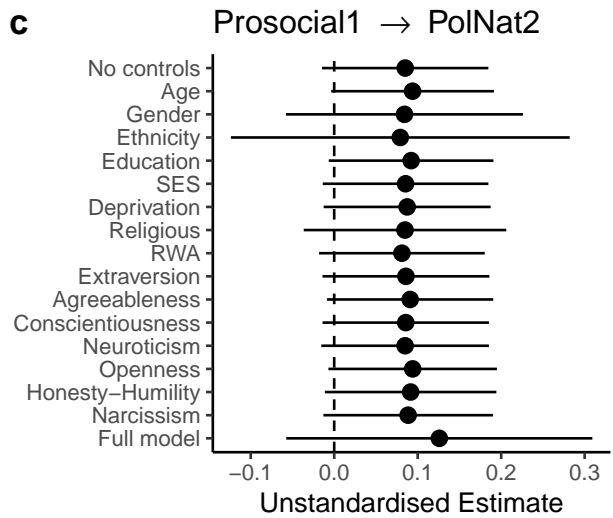**b** IncRed1 → Prosocial2   **c** Prosocial1 → IncRed2

Supplementary Figure 9. *Results of cross-lagged panel model with the prosocial phenotype and support for income redistribution, analysing listwise-deleted data.* (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Support for income redistribution is treated as ordinal. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.

a

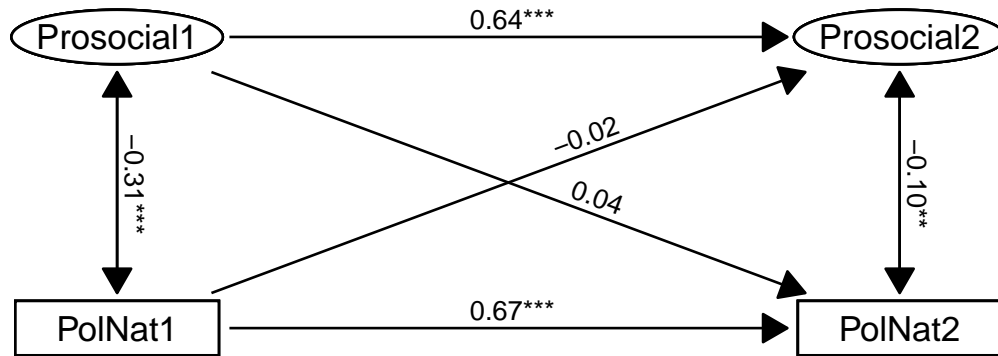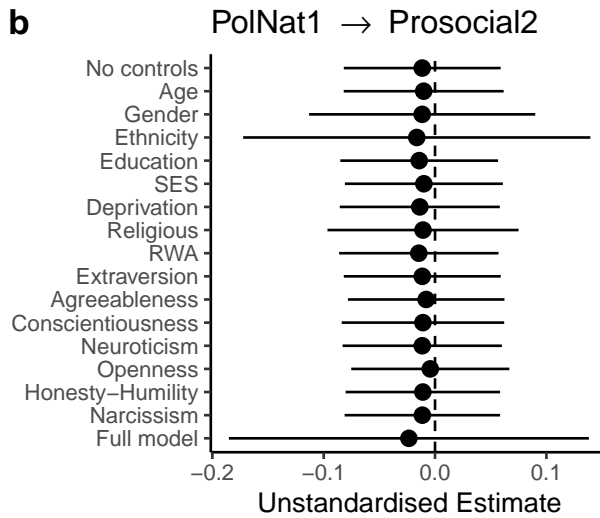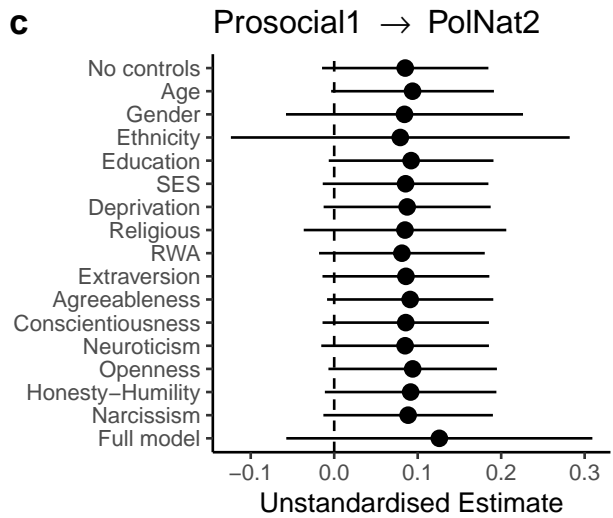**b** IncAtt1 → Prosocial2   **c** Prosocial1 → IncAtt2
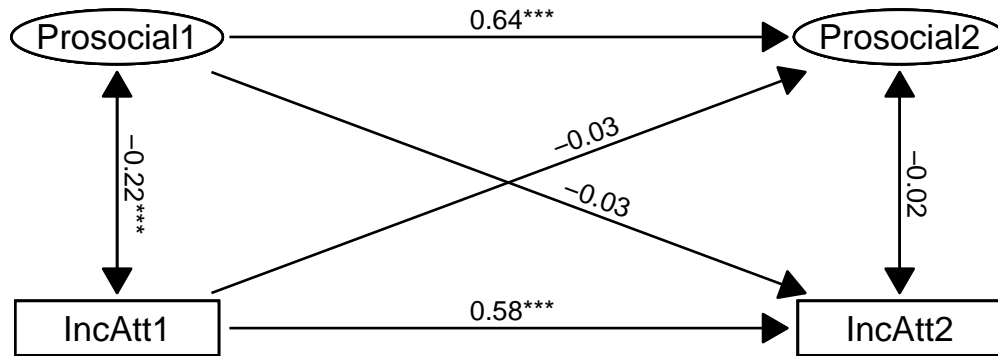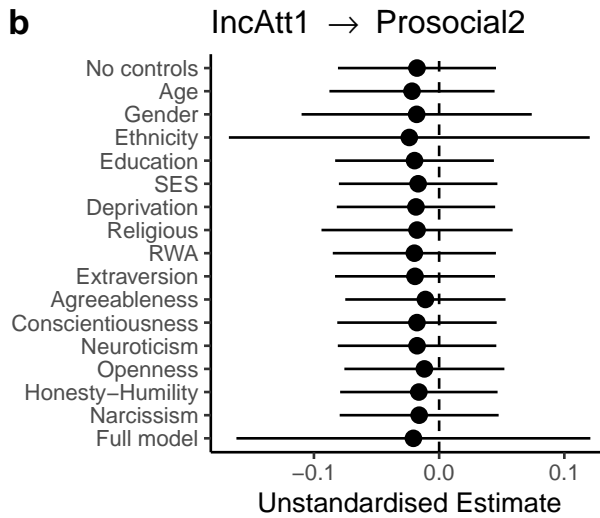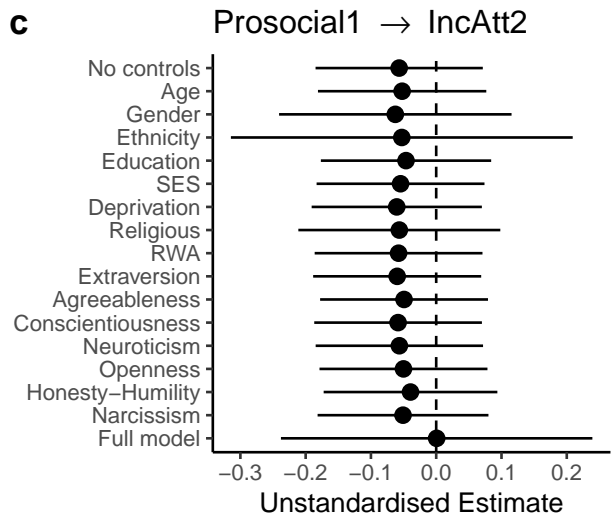
Supplementary Figure 10. *Results of cross-lagged panel model with the prosocial phenotype and income attribution beliefs, pooling over 20 imputed datasets.* (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Income attribution beliefs are treated as ordinal. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.

Supplementary Figure 11. *Results of cross-lagged panel model with the prosocial phenotype and income attribution beliefs, analysing listwise-deleted data.* (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Income attribution beliefs are treated as ordinal. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.

Supplementary Figure 12. Results of cross-lagged panel model with the prosocial phenotype and support for the National Party, pooling over 20 imputed datasets. (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Support for the National Party is treated as ordinal. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.

a

**b**   PolNat1 → Prosocial2   **c**   Prosocial1 → PolNat2

Supplementary Figure 13. *Results of cross-lagged panel model with the prosocial phenotype and support for the National Party, analysing listwise-deleted data.* (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Support for the National Party is treated as ordinal. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.
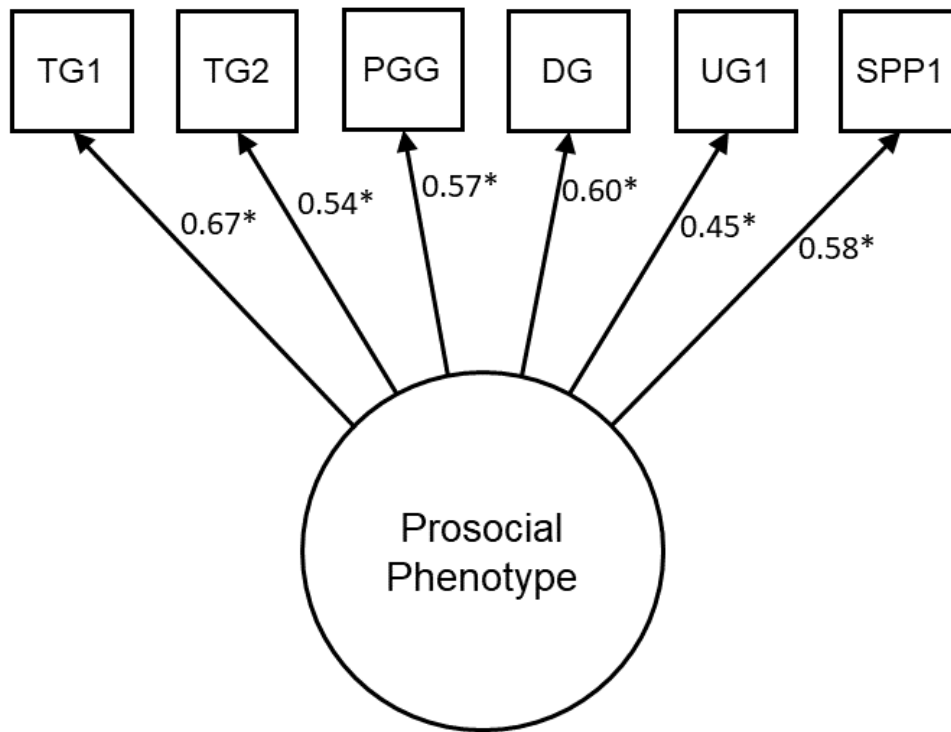
Supplementary Figure 14. *Confirmatory factor model for the prosocial phenotype in Wave 2, including additional games.* TG1 and SPP1 are treated as binary endogenous variables. Numbers are standardised coefficients. *$p < 0.05$. TG1 = Trust Game (Give), TG2 = Trust Game (Return), PGG = Public Goods Game, DG = Dictator Game, UG1 = Ultimatum Game (Offer), SPP1 = Second-Party Punishment Game.
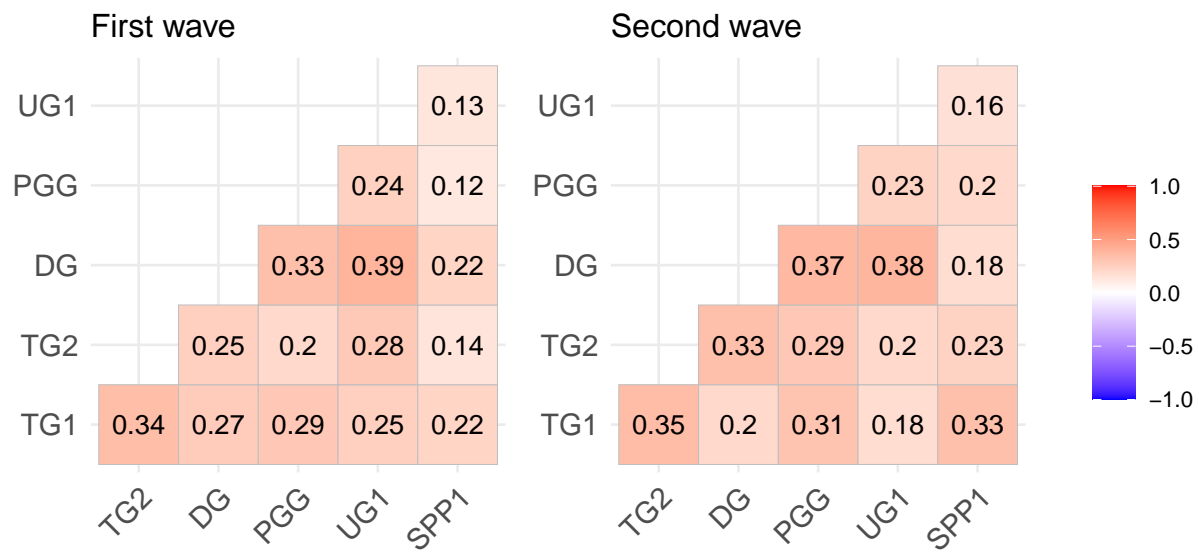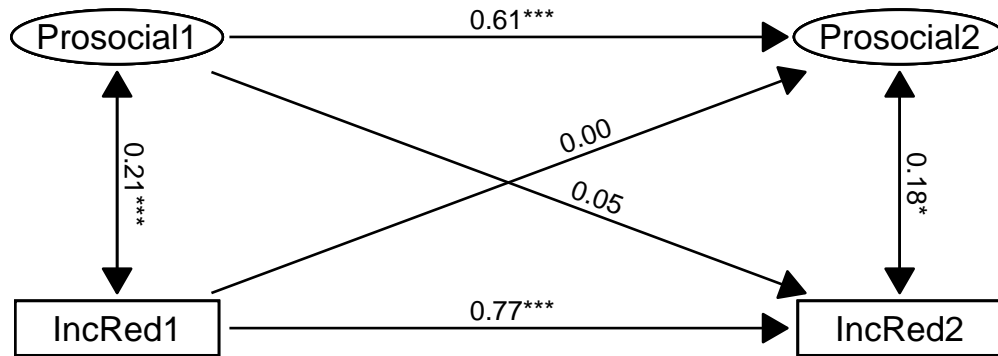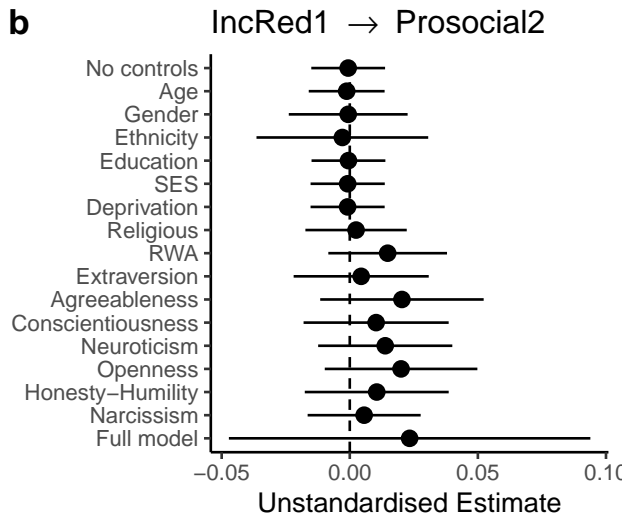
Supplementary Figure 15. *Zero-order correlations between game decisions in the first and second wave, including additional games.* All *p*-values < 0.01. TG1 = Trust Game (Give), TG2 = Trust Game (Return), DG = Dictator Game, PGG = Public Goods Game, UG1 = Ultimatum Game (Offer), SPP1 = Second-Party Punishment Game (Give).
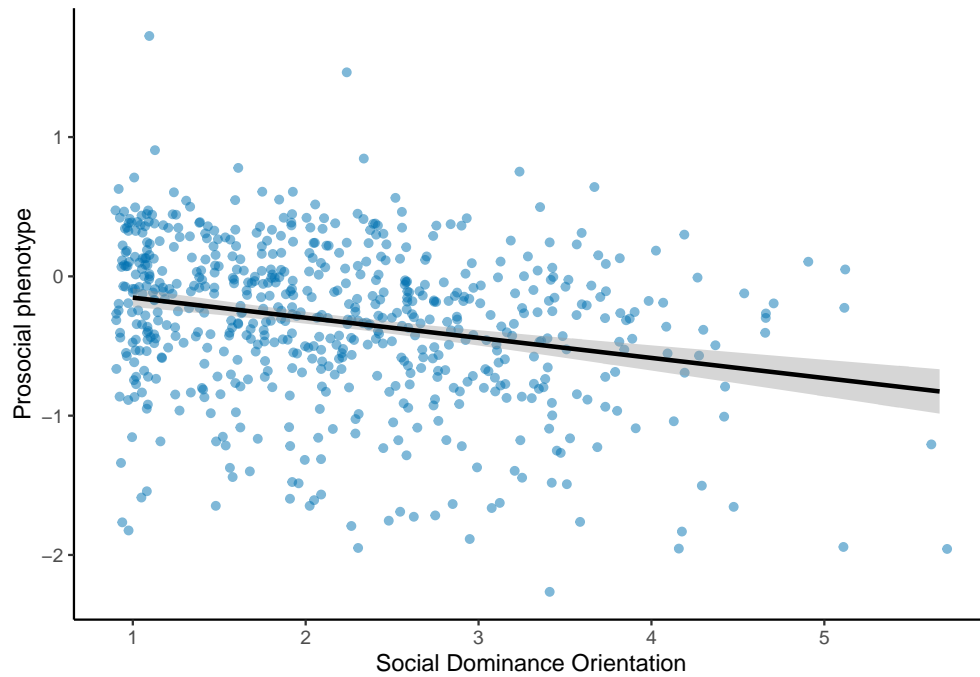
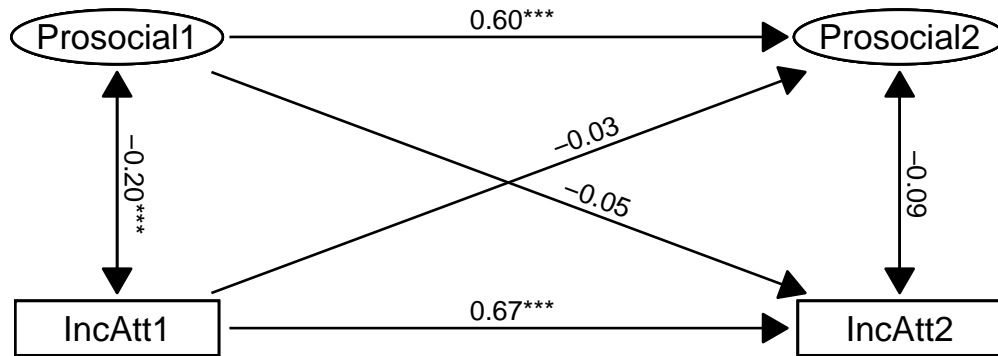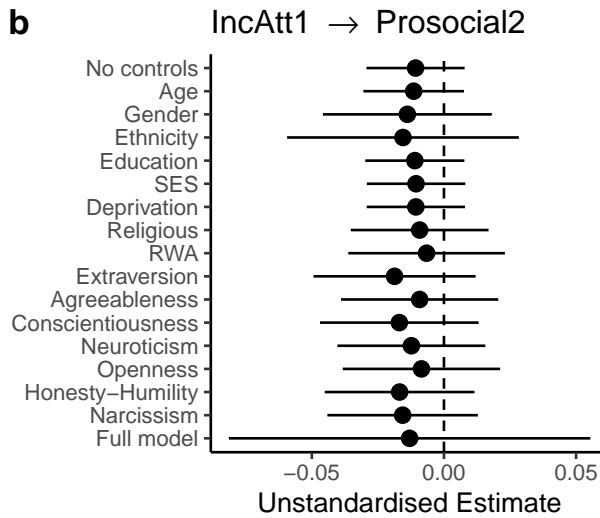*Supplementary Figure 16*. *Social Dominance Orientation (mean score) is negatively related to model-predicted prosocial latent variable scores in the second wave, when additional games (Ultimatum Game, Second-Party Punishment Game) are included in the measurement model.*
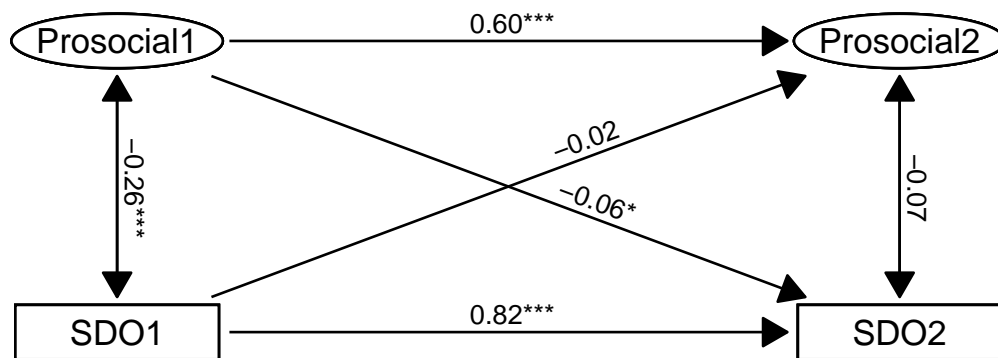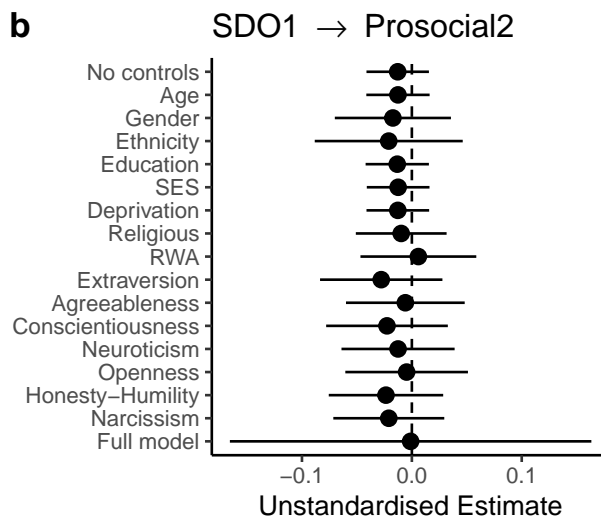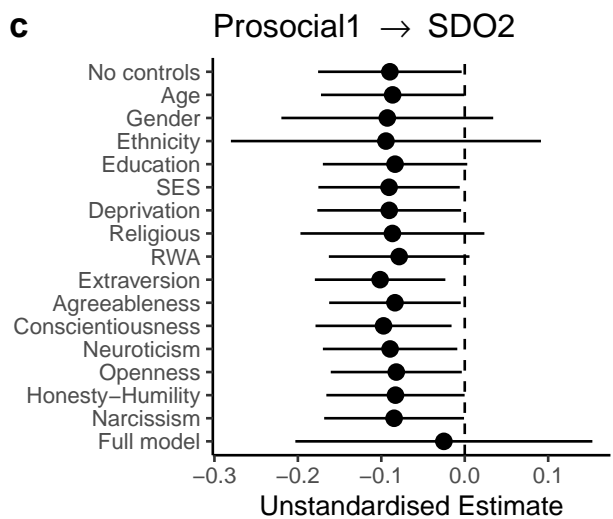
Supplementary Figure 17. Results of cross-lagged panel model with the prosocial phenotype and SDO, pooling over 20 imputed datasets and including the Ultimatum Game and the Second-Party Punishment Game in the measurement model for the prosocial phenotype. (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.

Supplementary Figure 18. *Results of cross-lagged panel model with the prosocial phenotype and support for income redistribution, pooling over 20 imputed datasets and including the Ultimatum Game and the Second-Party Punishment Game in the measurement model for the prosocial phenotype.* (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Support for income redistribution is treated as ordinal. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.

Supplementary Figure 19. Results of cross-lagged panel model with the prosocial phenotype and income attribution beliefs, pooling over 20 imputed datasets and including the Ultimatum Game and the Second-Party Punishment Game in the measurement model for the prosocial phenotype. (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Income attribution beliefs are treated as ordinal. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.

a

b PolNat1 → Prosocial2   **c** Prosocial1 → PolNat2

Supplementary Figure 20. Results of cross-lagged panel model with the prosocial phenotype and support for the National Party, pooling over 20 imputed datasets and including the Ultimatum Game and the Second-Party Punishment Game in the measurement model for the prosocial phenotype. (a) Autoregressive effects, cross-lagged effects, and within-wave (residual) correlations from the model. Support for the National Party is treated as ordinal. Note that measurement models for the prosocial phenotype latent variables are omitted from this figure. Numbers are standardised coefficients, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (b, c) Forest plots visualising the change in cross-lagged paths when controlling for time-invariant covariates, individually and in a full model. Points are unstandardised estimates, lines are 95% confidence intervals.

Supplementary Tables

Supplementary Table 1

Self-report items from the New Zealand Attitudes and Values Study.

Item	Description / Text	Wave
SDO1	It is OK if some groups have more of a chance in life than others	10 - 11
SDO2	Inferior groups should stay in their place	10 - 11
SDO3	To get ahead in life, it is sometimes okay to step on other groups	10 - 11
SDO4 (reversed)	We should have increased social equality	10 - 11
SDO5 (reversed)	It would be good if groups could be equal	10 - 11
SDO6 (reversed)	We should do what we can to equalise conditions for different groups	10 - 11
RWA1	It is always better to trust the judgment of the proper authorities in government and religion than to listen to the noisy rabble-rousers in our society who are trying to create doubt in people's minds	10
RWA2	It would be best for everyone if the proper authorities censored magazines so that people could not get their hands on trashy and disgusting material	10
RWA3	Our country will be destroyed some day if we do not smash the perversions eating away at our moral fibre and traditional beliefs	10
RWA4 (reversed)	People should pay less attention to The Bible and other old traditional forms of religious guidance, and instead develop their own personal standards of what is moral and immoral	10
RWA5 (reversed)	Atheists and others who have rebelled against established religions are no doubt every bit as good and virtuous as those who attend church regularly	10
RWA6 (reversed)	Some of the best people in our country are those who are challenging our government, criticizing religion, and ignoring the 'normal way' things are supposed to be done	10
Income redistribution	Redistributing money and wealth more evenly among a larger percentage of the people in New Zealand through heavy taxes on the rich	10 - 11
Income attribution	If incomes were more equal, people would be less motivated to work hard	10 - 11
Support for National Party	Level of support for The National Party	10 - 11
Age	What is your date of birth?	10
Gender	What is your gender? (open-ended)	10
Ethnicity	Which ethnic group do you belong to? (NZ census question)	10
Education level	NZ Reg (0-10 education ordinal rank)	10
Socio-economic status	NZSEI13 (NZ Socio-economic index)	10
Local deprivation	Deprivation score 2013 (for Meshblock)	10
Religiosity	Do you identify with a religion and/or spiritual group?	10
Extraversion1	Am the life of the party	10
Extraversion2 (reversed)	Don't talk a lot	10
Extraversion3 (reversed)	Keep in the background	10

Table 1 continued

Item	Description / Text	Wave
Extraversion4	Talk to a lot of different people at parties	10
Agreeableness1	Sympathize with others' feelings	10
Agreeableness2 (reversed)	Am not interested in other people's problems	10
Agreeableness3	Feel others' emotions	10
Agreeableness4 (reversed)	Am not really interested in others	10
Conscientiousness1	Get chores done right away	10
Conscientiousness2	Like order	10
Conscientiousness3 (reversed)	Make a mess of things	10
Conscientiousness4 (reversed)	Often forget to put things back in their proper place	10
Neuroticism1	Have frequent mood swings	10
Neuroticism2 (reversed)	Am relaxed most of the time	10
Neuroticism3	Get upset easily	10
Neuroticism4 (reversed)	Seldom feel blue	10
Openness1	Have a vivid imagination	10
Openness2 (reversed)	Have difficulty understanding abstract ideas	10
Openness3 (reversed)	Do not have a good imagination	10
Openness4 (reversed)	Am not interested in abstract ideas	10
Narcissism1	Feel entitled to more of everything	10
Narcissism2	Deserve more things in life	10
Honesty-Humility1 (reversed)	Would like to be seen driving around in a very expensive car	10
Honesty-Humility2 (reversed	Would get a lot of pleasure from owning expensive lucury goods	10

Supplementary Table 2

Results of multilevel models estimating the differences in game behaviour between the two waves. There are no significant differences between waves for the Trust Game and the Public Goods Game. In the Dictator Game, participants give 2 fewer points (out of 100) in the second wave, on average.

Game	b	SE	t	p
Trust Game (Give)	0.02	0.02	0.72	0.47
Trust Game (Return)	0.00	0.01	-0.42	0.67
Public Goods Game	0.02	0.01	1.28	0.20
Ultimatum Game (Offer)	0.00	0.01	0.01	0.99
Dictator Game	-0.02	0.01	-2.35	0.02
Second Party Punishment Game (Give)	-0.03	0.02	-1.93	0.05

Supplementary Table 3

Results of linear models estimating the differences in key variables between participants who dropped out (n = 414) and participants who completed both waves of data collection (n = 631).

Variable	b	SE	t	p
egame.TG1.T10	0.00	0.03	-0.05	0.96
egame.TG2.T10	-1.11	1.83	-0.61	0.55
egame.PGG.T10	1.43	1.99	0.72	0.47
egame.UG1.T10	-1.18	0.80	-1.48	0.14
egame.DG.T10	-0.31	1.26	-0.25	0.80
egame.SPP1.T10	-0.07	0.02	-3.22	0.00
T10.SDO	0.05	0.06	0.86	0.39
Issue.IncomeRedistribution.T10	0.03	0.12	0.28	0.78
IncomeAttribution.T10	0.04	0.12	0.38	0.70
Pol.SupNational.T10	0.15	0.13	1.20	0.23

Supplementary Table 4

Measurement invariance analysis of the prosocial phenotype latent variable supports strict measurement invariance. These models include the Trust Game (Give), Trust Game (Return), Public Goods Game, and Dictator Game in the measurement model for the prosocial phenotype. CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardised Root Mean Square Residual.

Model	χ^2	df	CFI	RMSEA	SRMR	ΔCFI	ΔRMSEA	ΔSRMR
Configural	26.70	15	0.991	0.035	0.038	-	-	-
Metric	37.61	18	0.985	0.042	0.044	-0.006	0.006	0.005
Scalar	41.90	22	0.985	0.038	0.044	0.000	-0.004	0.000
Strict	44.87	25	0.985	0.036	0.044	0.000	-0.002	0.001

Supplementary Table 5

Measurement invariance analysis of the prosocial phenotype latent variable supports strict measurement invariance. These models include the Trust Game (Give), Trust Game (Return), Public Goods Game, Dictator Game, Ultimatum Game, and Second-Party Punishment Game in the measurement model for the prosocial phenotype. CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardised Root Mean Square Residual.

Model	χ^2	df	CFI	RMSEA	SRMR	ΔCFI	ΔRMSEA	ΔSRMR
Configural	111.89	47	0.975	0.047	0.056	-	-	-
Metric	125.96	52	0.971	0.048	0.058	-0.004	0.001	0.002
Scalar	133.50	58	0.970	0.045	0.058	-0.001	-0.002	0.000
Strict	138.59	62	0.970	0.044	0.058	0.000	-0.001	0.000