European Journal of Personality, *Eur. J. Pers.* **32**: 116–127 (2018)
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/per.2149

Lead Us (Not) into Temptation: Testing the Motivational Mechanisms Linking Honesty–Humility to Cooperation

BENJAMIN E. HILBIG[1,2]*, PASCAL J. KIESLICH[3], FELIX HENNINGER[1,2,3], ISABEL THIELMANN[1] and INGO ZETTLER[4]

[1]*Cognitive Psychology Lab, University of Koblenz-Landau, Germany*
[2]*Max-Planck-Institute for Research on Collective Goods, Germany*
[3]*University of Mannheim, Germany*
[4]*University of Copenhagen, Denmark*

Abstract: Over the past decades, there has been considerable interest in individual differences in cooperative behaviour and how these can be explained. Whereas the Honesty–Humility dimension from the HEXACO model of personality has been identified as a consistent predictor of cooperation, the underlying motivational mechanisms of this association have remained unclear—especially given the confound between the temptation to exploit others and the fear of being exploited as motivational drivers of defection in social dilemmas. In a reanalysis and a new experiment, we tease apart these mechanisms by manipulating the rank order of pay-offs in a symmetric two-person game paradigm, essentially implementing the classic prisoner's dilemma, stag hunt, and chicken games. Results revealed that Honesty–Humility predicted cooperation specifically in the games in which temptation was a potential motivator of defection, whereas it did not account for cooperation in those games in which only fear implied defection. Our findings thereby shed light on the underlying motivational mechanisms of the Honesty–Humility–cooperation link and, more generally, demonstrate how economic games can be used to disentangle such mechanisms. Copyright © 2018 European Association of Personality Psychology

Key words: Honesty-Humility; motivation; HEXACO; social dilemmas; cooperation

INTRODUCTION

Cooperation and other forms of prosocial behaviour are vital pillars of a functioning society: They ensure efficient interactions and prevent conflict. The critical nature of cooperation is most apparent in so-called social dilemmas (Dawes, 1980; Kollock, 1998) in which individual utility maximization is at odds with collective efficiency: Whereas the largest individual outcomes are achieved through non-cooperation, free-riding, or exploiting others, these very behaviours—especially if pursued by many or most individuals—reduce social welfare and thus lead to diminished collective outcomes. Indeed, decades of research have revealed that a substantial proportion of individuals are willing to cooperate in social dilemmas (Colman, 2003; Sally, 1995; Zelmer, 2003) and that the evolution of cooperation hinges on both a willingness to refrain from exploitation and a willingness to forgive non-cooperators to a certain extent and reestablish cooperation (Axelrod & Hamilton, 1981; Nowak, 2006).

*Correspondence to: Benjamin E. Hilbig, Cognitive Psychology Lab, University of Koblenz and Landau, Fortstraße 7, 76829 Landau, Germany.
E-mail: hilbig@uni-landau.de

This article earned Open Data and Open Materials badges through Open Practices Disclosure from the Center for Open Science: https://osf.io/tvyxz/wiki. The data and materials are permanently and openly accessible at https://osf.io/w2xsr/. Author's disclosure form may also be found at the Supporting Information in the online version.

To understand why some individuals opt for non-exploitation and do so across varied situations and contexts, a growing body of research has investigated personality traits as explanatory variables. Across the diverse approaches ranging from relatively narrow, specific traits such as social values (Balliet, Parks, & Joireman, 2009; Van Lange, 1999; Van Lange, De Bruin, Otten, & Joireman, 1997) to broad, basic personality dimensions (for a meta-analytic review, see Zhao & Smillie, 2015), there is growing consensus that personality plays a substantial role in predicting cooperation. With particular consistency across studies, the Honesty–Humility dimension from the HEXACO model of personality structure (Ashton & Lee, 2007, 2008; Ashton, Lee, & De Vries, 2014) has been linked to cooperation in social dilemmas (Hilbig, Zettler, & Heydasch, 2012; Kieslich & Hilbig, 2014; Mischkowski & Glöckner, 2016; Ruch, Bruntsch, & Wagner, 2017; Zettler, Hilbig, & Heydasch, 2013) and other prosocial, non-exploitative behaviours (Ackermann, Fleiß, & Murphy, 2016; Hilbig, Glöckner, & Zettler, 2014; Hilbig, Thielmann, Hepp, Klein, & Zettler, 2015; Zhao, Ferguson, & Smillie, 2017).

On closer inspection, however, this state of knowledge is actually incomplete in that the motivational mechanisms that link Honesty–Humility to observable decisions in social dilemmas and economic games more generally are largely unknown. Stated bluntly, we have learned that a trait that is defined as 'the tendency to be fair and genuine in dealing

Handling editor: Mitja Back
Received 24 October 2017
Revised 2 March 2018, Accepted 7 March 2018

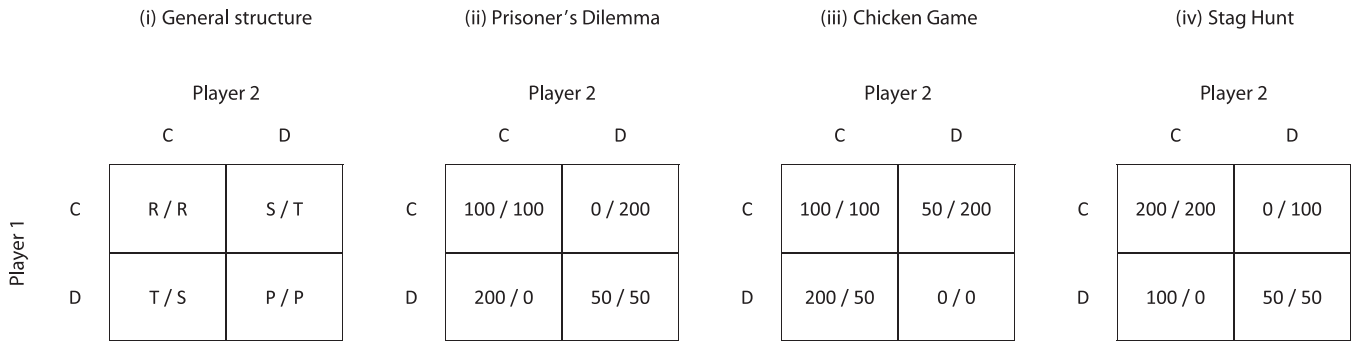

Figure 1. General pay-off structure (normal form representation) of a two-person simultaneous game (panel i) and exemplary pay-off matrices for a prisoner's dilemma (panel ii), a chicken game (panel iii), and a stag hunt game (panel iv). *C* and *D* stand for each player's choice options, that is, cooperation and defection, respectively. The first value in each cell represents the pay-off of player 1, and the second the pay-off of player 2.

with others, in the sense of *cooperating with others even when one might exploit them* without suffering retaliation' (Ashton & Lee, 2007, p. 156, emphasis added) and measured correspondingly (consider items such as 'I wouldn't cheat a person even if he or she was a real "sucker"'[1]) in the widely used HEXACO Personality Inventory-Revised (Ashton & Lee, 2009; Lee & Ashton, 2004, 2006) predicts whether individuals cooperate. This finding is important per se as it corroborates the very definition of Honesty–Humility and, more generally, in light of how rarely actual behaviour is studied in personality research (Baumeister, Vohs, & Funder, 2007; Funder, 2001; Funder, 2009). However, it is a surface-level association that cannot provide insight on the underlying causal mechanisms or motivational processes at work. Although this 'black-box' problem arguably applies to a substantial proportion of personality research, it is particularly unsatisfactory in the present case because the (cooperation) behaviour to which Honesty–Humility has been linked actually entails multiple possible motivations behind the same observable decision (Thielmann, Böhm, & Hilbig, 2015).

The issue of multiple motivations driving the same choice is most easily demonstrated by means of the classic prisoner's dilemma game (PDG; Rapoport & Chammah, 1965), a two-person simultaneous game in which each of two players independently faces a choice between two possible moves, cooperation (C) and defection (D). As shown in Figure 1, pay-offs are determined by the combination of players' choices such that each receives *reward* R for mutual cooperation (C, C), *punishment* P for mutual defection (D, D), *temptation* T for unilateral defection (D, C), and *sucker* S for unilateral cooperation (C, D). By definition, the pay-off structure of the PDG is $T > R > P > S$, and therefore, defection strictly dominates cooperation: Independent of the other player's move, one is always better off by defecting ($T > R$ and $P > S$), and thus, mutual defection is the game's unique Nash equilibrium. However, if both players defect, their pay-off is smaller than that of mutual cooperation ($R > P$). An exemplary PDG pay-off matrix is shown in panel ii of Figure 1.

Given this structure, the decision to defect can be an expression of two sets of motivations: One is essentially the *temptation* inherent in unilateral defection by which—in terms of Interdependence Theory (Kelley & Thibaut, 1978; Thibaut & Kelley, 1959)—one can maximize one's own pay-off (greed), minimize the other's pay-off (aggression), or maximize the difference between one's own and the other's pay-off (competitiveness). The other is the *fear* of winding up as the sucker, that is, cooperating unilaterally and being exploited by the other. Stated simply, defection can follow from one's own willingness to exploit the other, one's fear of being exploited in this very manner, or both (Ahn, Ostrom, Schmidt, Shupp, & Walker, 2001; Bruins, Liebrand, & Wilke, 1989).[2]

Although both motivations may drive cooperation behaviour in social dilemmas, prior research suggests that greed (i.e. temptation) is indeed the predominant driver of a (low) willingness to cooperate (Poppe & Utens, 1986; Rapoport & Eshed-Levy, 1989). For example, Poppe and Utens (1986) used three variants of a step-level public goods game in which they manipulated whether greed alone, fear alone, or both would drive defection and found that individuals contributed less when greed (alone) was involved than when fear (alone) was involved. However, the relative weight of motivations may also depend on the specific pay-off structure. Specifically, using a public goods game, Yamagishi and Sato (1986) manipulated whether the good provided to all group members was based on the lowest individual contribution (conjunctive condition) or on the highest individual contribution (disjunctive contribution) to the public good. In line with the idea that the former situation should trigger fear whereas the latter should trigger greed, individual levels of fear were negatively related to cooperation in the conjunctive condition but not in the disjunctive condition, whereas the opposite was true for individual levels of greed.

By implication, the finding that individuals low in Honesty–Humility are more likely to defect in social dilemmas can be attributed to at least two underlying motivational mechanisms that mirror the classical distinction between approach and avoidance goals in motivation (Atkinson, 1957; McClelland, 1951): one's pull towards (approach) versus desire to evade (avoidance) certain outcomes (Elliot, 1999; Elliot & Covington, 2001). First, pertaining to approach motivation,

[1]Note that the Honesty–Humility scale involves additional items that do not primarily pertain to fairness versus exploitation but to other aspects of this broad factor (such as modesty).

[2]Note that fear may not necessarily result from an inherent belief that one will be exploited, but it may also result from the belief that the other might experience fear (i.e. fear of another's fear).

DOI: 10.1002/per

it could be that individuals low in Honesty–Humility are particularly susceptible to the *temptation of exploiting others*, whereas their counterparts high in Honesty–Humility are less responsive to the tempting character of defection. Second, and related to avoidance motivation, it might be that individuals low in Honesty–Humility *fear being exploited by others*, whereas their counterparts high in Honesty–Humility are less sensitive to the risk inherent in cooperating.

Indeed, prior evidence does hint that the links between Honesty–Humility and decisions in economic games can be attributed to temptation or greed, respectively. For one, Honesty–Humility has been consistently linked to dictator game (Forsythe, Horowitz, Savin, & Sefton, 1994) giving (Hilbig et al., 2015; Zhao & Smillie, 2015). In this game, the participant simply decides how to allocate a resource between herself or himself and another individual who cannot respond in any way. Clearly, fear cannot be a driver of behaviour in this game (as the recipient is completely powerless), and thus, the positive link between Honesty–Humility and giving indicates that individuals low in Honesty–Humility are motivated by greed, aggression, or competitiveness. Moreover, one prior study using a hypothetical PDG manipulated the—explicitly stated—probability of the second player cooperating and found that Honesty–Humility was only associated with cooperation if the other player was relatively likely to cooperate (Zettler et al., 2013, Exp. 3). Thus, Honesty–Humility predicted cooperation only if the other player could be exploited (temptation). However, as the study used a hypothetical game (and thus no actual temptation) and explicitly stated cooperation probabilities (rendering the game asymmetric), the findings can only serve as a first toehold for the temptation mechanism.

In this regard, it is also important to note that, although temptation appears to be more clearly implied by the definition of Honesty–Humility and corresponding evidence (sketched above), fear might actually be a result of one's own temptation: If individuals low in Honesty–Humility are particularly receptive to temptation, this should increase their fear of being exploited. It has been argued and shown that—in the face of not knowing whether to expect cooperation—people will 'project' their own dispositions onto others (Krueger, 2007; Krueger & Acevedo, 2007), which has already been demonstrated for Honesty–Humility in a trust game (Thielmann & Hilbig, 2014). Specifically, it was found that individuals base their trustworthiness expectations (whether they expect others to respond in a fair rather than exploitative manner to trust) on their own trustworthiness, that is, their own tendency to cooperate. So if those low in Honesty–Humility project their own tendency to be tempted by the benefits of defection onto others, they must also expect others to take advantage of them. To avoid being exploited, they themselves must not cooperate. Thereby, both mechanisms—temptation and fear—may work in combination. The main goal of the present work is to uncover whether either or both of these mechanisms provide a viable explanation of the relationship between Honesty–Humility and cooperation.

Fortunately, on the basis of straightforward game-theoretic considerations, it is possible to dissect the motivational processes in question. Specifically, variation of the game's pay-off structure yields game variants in which exactly one of the motivations under scrutiny leads to distinct choices as compared with the classic PDG. First, in the so-called chicken game (de Heus, Hoogervorst, & Dijk, 2010; Rapoport & Chammah, 1969), the rank order of pay-offs S and P is reversed (i.e. $T > R > S > P$) such that the worst outcome no longer occurs for unilateral cooperation but actually if both players defect. An exemplary chicken game pay-off matrix is shown in panel iii of Figure 1. As can be seen therein, defection still is tempting because unilateral defection produces the largest pay-off (200), but cooperation is no longer risky because being exploited (unilateral cooperation) is better (50) than mutual defection (0). As a consequence, fear no longer motivates defection (because $S > P$ unlike in the PDG) whereas temptation still motivates defection (because $T > R$ as in the PDG).

Second, in the assurance or stag hunt game (Gächter, 2004), the rank order of pay-offs T and R is reversed (i.e. $R > T \geq P > S$) such that the best outcome is no longer associated with unilateral defection but actually associated with mutual cooperation. An exemplary stag hunt pay-off matrix is shown in panel iv of Figure 1. As can be seen therein, defection is no longer tempting because the highest pay-off can be gained through mutual cooperation (200), but cooperation is still risky because unilateral cooperation produces a lower pay-off (0) than does mutual defection (50). Consequently, temptation no longer motivates defection (because $R > T$ unlike in the PDG), but fear does (because $P > S$ as in the PDG).

Importantly, neither the chicken game nor the stag hunt game can be taken to provide a pure (isolated) measure of temptation or fear, respectively. However, for each game in comparison with the PDG, exactly one of these motivations no longer implies defection: In the PDG, there is one Nash equilibrium (D, D), implying that one is always better off to defect, no matter what one expects the other player to do (independent of whether one intends to exploit her or him [i.e. temptation] or protect oneself against exploitation [i.e. fear]). In the chicken game, there are two Nash equilibria (D, C) and (C, D), implying that one is not always better off to defect, but only if the other cooperates. Thus, if one expects the other to cooperate, one should defect (exploiting the other, i.e. temptation), but if one expects the other to defect, one should cooperate (thus, fear of exploitation does not imply defection, but cooperation). In turn, in the stag hunt game, there are also two Nash equilibria (C, C) and (D, D), again implying that one is not always better off to defect, but only if the other defects. Thus, if one expects the other to defect, one should defect (avoiding the sucker pay-off, i.e. fear), but if one expects the other to cooperate, one should cooperate (thus, temptation does not imply defection, but cooperation).

Taken together, a comparison across these games (i.e. experimentally manipulating the rank-order of pay-offs) can yield insights into which of the motivations in question—temptation versus fear—links Honesty–Humility to cooperation: If temptation (alone) is the crucial factor, then Honesty–Humility should positively predict cooperation in the PDG and the chicken game but *not* in the stag hunt game. If fear (alone) is responsible, Honesty–Humility should be positively linked to cooperation in the PDG and the stag hunt

DOI: 10.1002/per

game, but *not* in the chicken game. Finally, if both mechanisms are at work, Honesty–Humility should predict cooperation in every game variant, but more so in the PDG (in which both mechanisms imply defection) than in the chicken game and stag hunt game (in each of which only one mechanism implies defection).

REANALYSIS

As a first test-bed for these competing predictions, we performed a series of new analyses on data from an earlier study (Kieslich & Hilbig, 2014) in which participants made decisions in all three games described earlier. Whereas the original study focused on the analysis of response dynamics—specifically the mouse trajectories recorded during individuals' decisions—we herein focus on the personality and choice data to provide a first test of the motivational processes underlying the link between Honesty–Humility and cooperation. A short overview of the methods employed in the study is given in the following; a full description is available in the original paper (Kieslich & Hilbig, 2014).[3]

Methods

The study consisted of two parts. First, participants completed a brief demographic survey and the 60-item version of the HEXACO Personality Inventory-Revised (HEXACO-60; Ashton & Lee, 2009; also http://hexaco.org) in its German translation (for psychometric properties, see Moshagen, Hilbig, & Zettler, 2014). Participants completed this prequestionnaire either online at least 48 hours prior to the study or in the laboratory directly preceding the study. The Honesty–Humility scale, on which we focus in the following, yielded acceptable internal consistency and typical descriptives (Cronbach's $\alpha = .78$, $M = 3.32$, $SD = 0.64$).

In the second (lab-based) part of the study, participants provided informed consent and were then thoroughly introduced to the games. In particular, participants received information regarding the structure of the games (which were presented as pay-off matrices; Figure 1), their interaction partners (in each trial an unknown, randomly assigned individual who participated in the study simultaneously), and the pay-off scheme (the monetary incentives they could gain during the task). No deception was used at any point in the study, and choices were incentivized.

The experiment was implemented in OpenSesame (Mathôt, Schreij, & Theeuwes, 2012) using the psynteract plug-in for interaction between participants (Henninger, Kieslich, & Hilbig, 2017) and the mousetrap plug-in for recording mouse movements (Kieslich & Henninger, 2017). As mentioned earlier, the games were presented as pay-off matrices, with incentives being expressed in points (and later transferred to monetary payouts). Both choice options for either player were labelled neutrally ('option A' and 'option B') so as to prevent socially desirable responding (Moshagen, Hilbig, & Musch, 2011), and the horizontal position of the cooperative option (left or right) was counterbalanced across participants.

The experiment consisted of 15 game trials in total with five trials per game variant, that is, PDG, chicken game, and stag hunt game, respectively (for the specific pay-off matrices used, refer to https://osf.io/w2xsr/ and Kieslich & Hilbig, 2014, Appendix A). In each trial anew, participants were randomly assigned to an interaction partner in the room. Participants received no feedback regarding their partners' decisions (and their corresponding pay-offs) over the course of the experiment. Participants' pay-offs were determined by the combination of their own and their interaction partners' choices in five randomly chosen trials. For these trials, points gained were converted into monetary payouts using a conversion rate of 0.50€ per 100 points.

Participants were recruited via a local participant panel at the University of Mannheim, Germany, and $N = 116$ individuals (79 female) took part in the study.[4] The mean age was 22.5 years (range 18 to 36, $SD = 4.0$). Participants were largely students from various disciplines. The materials, data, and analyses for the current reanalysis are available online at https://osf.io/w2xsr/.

Results and discussion

Descriptives of all variables, cooperation rates for each of the games, and their association with the HEXACO dimensions are summarized in Table 1. Across all games, participants cooperated in 58% of trials. Cooperation rates were lowest in the PDG (25%), followed by the chicken game (63%) and the stag hunt game (86%). To assess the predictive power of Honesty–Humility[5] for the willingness to cooperate, we first ran a generalized linear model with a binomial link function, predicting the probability of choosing the cooperative option on the basis of the individual level of Honesty–Humility for each game separately. The models were estimated using the glmer function of the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2017) and included a subject-level random intercept. In the PDG and the chicken game, the probability of cooperation significantly increased with increasing levels of Honesty–Humility, $OR = 2.39$, $z = 2.35$, $p = .019$, and $OR = 2.52$, $z = 3.72$, $p < .001$, respectively, both of which can be considered medium-sized effects (throughout all analyses, we interpret the odds ratios as effect sizes as per Rosenthal, 1996). By contrast, Honesty–Humility did not predict cooperation in the stag hunt game, $OR = 1.46$, $z = 0.85$, $p = .397$ (refer to Table 1 for corresponding zero-order correlations with cooperation rates).

In order to test the competing hypotheses outlined earlier, we further ran a generalized linear model in which we included Honesty–Humility (centred on the sample mean),

[3]The original paper is published as open access and can be obtained from http://journal.sjdm.org/14/14808/jdm14808.pdf. The study mainly analysed participants' mouse movements during the game decisions. An overview of different measures that can be derived from mouse movements and that were analysed in the original article is provided in Kieslich and Hilbig (2014, Appendix B).

[4]As in the original article, the following analyses are based on 115 participants excluding the one participant who always cooperated. Nonetheless, results were replicated for all 116 participants.

[5]For the present reanalysis and the following experiment, we also provide supplementary analyses on the facet and item levels, which are available online at https://osf.io/w2xsr/.

DOI: 10.1002/per

Table 1. Means, standard deviations (in parentheses), and intercorrelations (with 95% confidence intervals in brackets) between the HEXACO dimensions and game behaviour (% cooperative decisions in each game) in the reanalysis

Variable	Range	*M* (*SD*)	Correlations 1	2	3	4	5	6	7	8
1. Honesty–Humility	1–5	3.32 (0.64)	*.78*							
2. Emotionality	1–5	3.28 (0.63)	.09 [−.09, .27]	*.81*						
3. Extraversion	1–5	3.45 (0.65)	.05 [−.13, .23]	−.12 [−.30, .06]	*.84*					
4. Agreeableness	1–5	3.07 (0.53)	.33*** [.16, .49]	−.18* [−.36, .00]	.16 [−.02, .33]	*.71*				
5. Conscientiousness	1–5	3.54 (0.57)	.18 [−.01, .35]	.18* [.00, .36]	.18 [.00, .35]	−.01 [−.19, .18]	*.79*			
6. Openness	1–5	3.65 (0.56)	.11 [−.07, .29]	−.05 [−.23, .14]	.25** [.07, .42]	.15 [−.03, .33]	.02 [−.17, .20]	*.73*		
7. Chicken game	0–1	0.63 (0.31)	.33*** [.16, .49]	.17 [−.01, .35]	−.04 [−.23, .14]	.17 [−.01, .35]	.05 [−.13, .23]	.05 [−.13, .23]	—	
8. Prisoner's dilemma game	0–1	0.25 (0.31)	.19* [.01, .36]	.12 [−.07, .29]	.12 [−.07, .30]	.06 [−.13, .24]	.13 [−.06, .30]	.21* [.03, .38]	.21* [.03, .38]	—
9. Stag hunt game	0–1	0.86 (0.24)	.10 [−.08, .28]	−.06 [−.24, .12]	.00 [−.18, .19]	−.04 [−.22, .15]	.02 [−.16, .20]	.08 [−.10, .26]	−.08 [−.26, .10]	.23* [.05, .39]

Note: Internal consistency reliabilities (Cronbach's α) are displayed in the diagonal (in italics). $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

the type of game, and their interaction as concurrent predictors as well as a subject-level random intercept. The three-level factor game type (PDG vs. chicken vs. stag hunt) was split into two distinct contrast variables. The first contrasted the two motivational mechanisms by differentiating between the chicken game (in which only temptation should motivate defection) and the stag hunt game (in which only fear should motivate defection; i.e. PDG = 0; chicken game = +1/2; stag hunt game = −1/2). The second contrast variable additionally captured the difference between the PDG (in which both fear and temptation may motivate defection) and the chicken and stag hunt game (i.e. PDG = −2/3; chicken game = +1/3; stag hunt game = +1/3). Finally, Honesty–Humility and the game type were permitted to interact, which constitutes the critical test of a differential influence of Honesty–Humility depending on the game—and corresponding motivational mechanism(s).

Corresponding to the observed cooperation rates across games, the regression model reflected a lower probability of cooperation in the chicken game than in the stag hunt game, $OR = 0.24$, $z = -8.84$, $p < .001$, and a higher probability of cooperation in both of these games than in the PDG, $OR = 15.47$, $z = 18.50$, $p < .001$. Both of these effect sizes can be considered large. Furthermore, the probability of cooperation generally increased with individuals' Honesty–Humility levels, albeit only resembling a small effect size, $OR = 1.78$, $z = 3.41$, $p < .001$. Most importantly, the influence of Honesty–Humility varied across games: In line with the analyses per game and the predictions of the temptation mechanism, Honesty–Humility had a stronger influence on the probability of cooperation in the chicken game than in the stag hunt game, $OR = 1.67$, $z = 2.09$, $p = .037$ (a small effect). By contrast, no significant difference in predictive power of Honesty–Humility emerged between the PDG on the one hand, and the chicken and stag hunt game on the other, $OR = 0.92$, $z = -0.39$, $p = .693$. The results of the model are visualized in Figure 2, which was created using the sjPlot package in R (Fox, 2003; Lüdecke, 2017).

In summary, the findings from the present reanalysis support the temptation mechanism, in that Honesty–Humility predicted cooperation in the PDG and the chicken game but not in the stag hunt game. In other words, Honesty–Humility only predicted cooperation in those games in which temptation motivates defection but not in the game in which only fear does so. This pattern was also confirmed when directly contrasting the predictive power of Honesty–Humility for cooperation in the chicken game versus stag hunt game, which showed a significantly stronger influence of Honesty–Humility in the former. In addition to these primary results, note that the positive main effect of Honesty–Humility on cooperation (across games) replicates previous findings (e.g. Zhao & Smillie, 2015), as does the main effect of the games themselves—with chicken and stag hunt game yielding substantially larger cooperation rates than the PDG (Butler, Burbank, & Chisholm, 2011; Haesevoets, Folmer, & Van Hiel, 2015; Rusch & Luetge, 2016). On the basis of these preliminary findings from our reanalysis, we conducted a new experiment to test whether temptation—rather than fear—is indeed the motivational driver of the link between Honesty–Humility and cooperation.

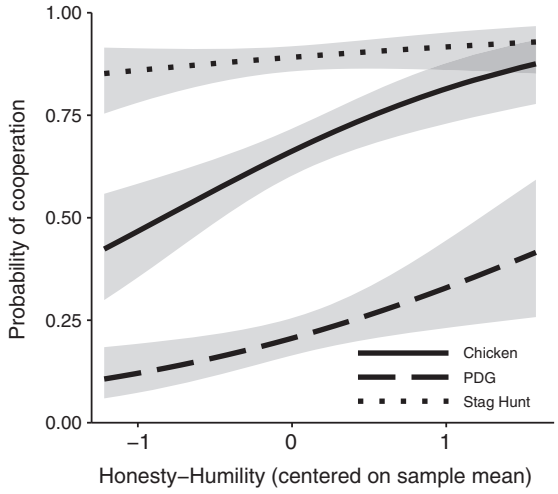

Figure 2. Probability of cooperation depending on Honesty–Humility (centred on the sample mean) and the type of game (reanalysis). Confidence bands indicate the 95% confidence intervals.

EXPERIMENT

As our main test of the temptation mechanism and aiming to replicate and slightly extend the reanalysis reported earlier, we conducted a novel experiment. First and foremost, we sought to conduct a conceptual replication of the finding that Honesty–Humility would predict cooperation in the PDG and chicken game, but not in the stag hunt game. In addition, aiming for a still more fine-grained test, we extended the experimental design to two variants of the PDG in addition to the chicken game and the stag hunt game. The two PDG variants differed in how cooperation-friendly the specific pay-off matrices were (Glöckner & Hilbig, 2012), an attribute that is well captured by the cooperation index (CI; Rapoport & Chammah, 1965; Vlaev & Chater, 2006). This index is defined as $CI = \frac{(R-P)}{(T-S)}$ and ranges from 0 to 1. Smaller values indicate that, relatively speaking, defection is very tempting and/or cooperation very risky (a cooperation-unfriendly matrix) whereas larger values indicate the opposite (a cooperation-friendly matrix). For the present experiment, we used a cooperation-unfriendly PDG with $CI = 0.2$ (similar to the PDG with $CI = 0.25$ in the reanalysis reported earlier) and a cooperation-friendly PDG with $CI = 0.8$ (Zettler et al., 2013). As the former PDG makes defection much more tempting than the latter, the to-be-tested temptation mechanism would predict a substantially larger positive influence of Honesty–Humility on cooperative behaviour in the PDG with $CI = 0.2$ than in the PDG with $CI = 0.8$. Across all four games, the temptation mechanism thus predicts a highly specific pattern: a positive relation between Honesty–Humility and cooperation in the PDG with $CI = 0.2$ and in the chicken game (games in which defection is tempting), but a much smaller or no relation between Honesty–Humility and cooperation in the PDG with $CI = 0.8$ and in the stag hunt game (games in which there is little to no temptation to defect).

Methods

The study consisted of two parts. In the first part—which participants completed either as an online pre-study or prior to the second part of the study in the lab—participants provided demographic information and filled in the German version of the HEXACO-60 (Ashton & Lee, 2009; Moshagen et al., 2014). As in Kieslich and Hilbig (2014), the Honesty–Humility scale yielded satisfactory internal consistency (Cronbach's $\alpha = .75$) and typical descriptives ($M = 3.27$, $SD = 0.62$; Moshagen et al., 2014).

The second part of the experiment was conducted in lab sessions of four to 10 participants and implemented with the software z-tree (Fischbacher, 2007). Upon arrival in the laboratory, participants first provided informed consent and received detailed information regarding the game. The task consisted of four trials in total, one for each game variant (i.e. PDG with $CI = 0.2$, PDG with $CI = 0.8$, chicken game, and stag hunt game), which were presented in random order. The games were shown as pay-off matrices, with pay-offs expressed as points and later transferred to monetary payouts. Choice options were labelled neutrally (option A and option B) so as to prevent socially desirable responding; the order of options was held constant (the first always corresponded to the cooperative choice). In each trial, participants were randomly assigned to another unknown participant in the room referred to as 'the other'. Monetary payouts were calculated on the basis of participants' own and their interaction partners' decisions in the four trials, with a conversion rate of 100 points = 0.50€. No feedback regarding the interaction partners' choices was provided during the experiment. After completion of the game trials, participants worked on additional unrelated tasks.

Participants were recruited from a local participant panel at the University of Mannheim, Germany. A total of $N = 260$ individuals completed the study. This sample size yields a highly satisfactory level of statistical power (as computed using G*Power, Faul, Erdfelder, Buchner, & Lang, 2009) of $1 - \beta = .96$ assuming an α level of .05, an effect of Honesty–Humility on cooperation of $OR = 2.39$ (as found for the PDG in the previous separate analysis with five choices per participant), and the analytic approach used herein (simple logistic regression with one choice per participant and game). Besides, it still implies a satisfactory power of $1 - \beta > .80$ for any effect that is at least small to medium sized ($OR = 2.0$; Rosenthal, 1996). Approximately two-thirds of participants (65.4%) were female; they were aged between 18 and 45 years ($M = 21.7$, $SD = 3.3$) and were students from diverse fields of study. The materials, data, and analyses of the current experiment are available at https://osf.io/w2xsr/.

Results and discussion

Table 2 provides descriptive statistics and correlations between the HEXACO dimensions and cooperation decisions in the experiment. Overall, participants cooperated in 65% of their choices across the four games. As in the previous reanalysis, the stag hunt game yielded the highest cooperation

DOI: 10.1002/per

Table 2. Means, standard deviations (in parentheses), and intercorrelations (with 95% confidence intervals in brackets) between the HEXACO dimensions and game behaviour (cooperative decision yes = 1, no = 0) as assessed in the experiment

			Correlations								
Variable	Range	*M* (*SD*)	1	2	3	4	5	6	7	8	9
1. Honesty–Humility	1–5	3.27 (0.62)	*.75*								
2. Emotionality	1–5	3.22 (0.65)	.03 [−.09, .15]	*.81*							
3. Extraversion	1–5	3.60 (0.59)	.03 [−.09, .15]	−.26*** [−.37, −.14]	*.80*						
4. Agreeableness	1–5	3.11 (0.52)	.20** [.08, .31]	.03 [−.09, .15]	.09 [−.03, .21]	*.72*					
5. Conscientiousness	1–5	3.58 (0.64)	.04 [−.08, .16]	.13* [.00, .24]	.16* [.04, .27]	.07 [−.05, .19]	*.84*				
6. Openness	1–5	3.60 (0.60)	.10 [−.02, .22]	−.13* [−.25, −.01]	.19** [.07, .31]	−.01 [−.13, .11]	.08 [−.04, .20]	*.74*			
7. Chicken game	0–1	0.70 (0.46)	.13* [.01, .25]	.13* [.01, .25]	−.08 [−.20, .04]	.01 [−.12, .13]	−.03 [−.15, .09]	.09 [−.03, .21]	—		
8. Prisoner's dilemma game (*CI* = 0.2)	0–1	0.39 (0.49)	.13* [.01, .25]	.01 [−.12, .13]	.02 [−.10, .14]	.05 [−.07, .18]	.00 [−.12, .12]	.05 [−.07, .17]	.15* [.03, .27]	—	
9. Prisoner's dilemma game (*CI* = 0.8)	0–1	0.68 (0.47)	.04 [−.08, .16]	−.04 [−.16, .08]	.03 [−.09, .15]	.06 [−.06, .18]	.08 [−.04, .20]	.01 [−.11, .13]	.15* [.03, .27]	.26*** [.15, .37]	—
10. Stag hunt game	0–1	0.82 (0.39)	−.02 [−.14, .10]	−.17** [−.28, −.05]	.09 [−.04, .21]	−.06 [−.18, .07]	−.05 [−.17, .08]	−.03 [−.15, .09]	−.07 [−.19, .05]	.09 [−.03, .21]	.21*** [.09, .32]

Note: Internal consistency reliabilities (Cronbach's α) are displayed in the diagonal (in italics). *$p < .05$; **$p < .01$; ***$p < .001$.

DOI: 10.1002/per

rate with 82%, followed by the chicken game with 70%, the cooperation-friendly PDG (CI = 0.8) with 68%, and the cooperation-unfriendly PDG (CI = 0.2) with 39%.

To examine the relation between Honesty–Humility and cooperation, we first ran separate generalized linear models for each game, predicting the probability of cooperation on the basis of the participants' individual Honesty–Humility score. The probability of cooperation significantly increased with increasing levels of Honesty–Humility in the PDG with CI = 0.2, OR = 1.56, z = 2.07, p = .039, and in the chicken game, OR = 1.57, z = 2.03, p = .042, with small effect sizes in both cases. By contrast, Honesty–Humility did not predict cooperation in the PDG with CI = 0.8, OR = 1.16, z = 0.69, p = .492, or in the stag hunt game, OR = 0.91, z = −0.36, p = .719 (refer to Table 2 for corresponding zero-order correlations).

To test the temptation mechanism across all games, we additionally ran a generalized linear model, predicting the probability of choosing the cooperative option on the basis of Honesty–Humility (centred on the sample mean), the type of game, and their interaction, and including a subject-level random intercept. Having added another PDG, we split the (now four-level) factor game into three distinct contrasts: To specifically test the temptation mechanism, the first contrast captured the difference between the two games in which defection was tempting (the PDG with CI = 0.2 and the chicken game; both coded with +1/4) and the two in which defection was far less tempting or not tempting at all (the PDG with CI = 0.8 and the stag hunt game; both coded with −1/4). The remaining contrasts were constructed orthogonally and differentiated the chicken game from the cooperation-unfriendly PDG with CI = 0.2 (i.e. chicken game = +1/2; PDG with CI = 0.2 = −1/2; otherwise 0) and the stag hunt game from the cooperation-friendly PDG (i.e. stag hunt game = +1/2; PDG with CI = 0.8 = −1/2; otherwise 0). As such, these contrasts are also particularly suited to test the fear mechanism: If fear drives the positive relation between Honesty–Humility and cooperation, Honesty–Humility should interact with both of these contrast variables given that each incorporates exactly one game variant with a high risk attached to cooperation (i.e. PDG with CI = 0.2 and stag hunt game, respectively) and one with a low risk attached to cooperation (i.e. chicken game and PDG with CI = 0.8, respectively).

As suggested by the cooperation rates per game, the regression analysis yielded significant main effects of the contrast variables, implying that the probability of cooperating was (i) lower in games with a high temptation to defect (i.e. PDG with CI = 0.2 and chicken game) than in games with a low temptation to defect (i.e. PDG with CI = 0.8 and stag hunt game), OR = 0.11, z = −6.98, $p < .001$ (large effect size), (ii) higher in the chicken game than in the PDG with CI = 0.2, OR = 4.67, z = 7.25, $p < .001$ (large effect size), and (iii) higher in the stag hunt game than in the PDG with CI = 0.8, OR = 2.40, z = 3.91, $p < .001$ (medium-sized effect). Regarding the influence of dispositional Honesty–Humility, analyses suggested a positive, albeit non-significant main effect of Honesty–Humility on cooperation that fell below the convention even for a small effect size, OR = 1.32, z = 1.80, p = .072; note, however, that a one-sided test of this directional hypothesis would imply statistical significance.

More importantly, and further supporting the temptation mechanism, there was a significant interaction between Honesty–Humility and the game type variable contrasting the games with high versus low temptation. Specifically, as depicted in Figure 3 and in line with the analyses per game, Honesty–Humility was a stronger predictor of cooperation in the chicken game and the cooperation-unfriendly PDG (CI = 0.2) than in the stag hunt game and the cooperation-friendly PDG (CI = 0.8), OR = 2.67, z = 1.99, p = .047 (medium-sized effect). For the two remaining contrast variables, no significant interactions with Honesty–Humility emerged, thus once more speaking against the fear mechanism. Specifically, there was neither a difference in the predictive power of Honesty–Humility for cooperation between the two games with a high temptation to defect (chicken game vs. PDG with CI = 0.2), OR = 0.99, z = −0.04, p = .965, nor between the two games with a low temptation to defect (stag hunt game vs. PDG with CI = 0.8), OR = 0.76, z = −0.77, p = .443. Overall, results provided further support that the relationship between Honesty–Humility and cooperation is attributable to individual differences in the sensitivity to temptation associated with defection, rather than to fear associated with cooperation.

On a more exploratory level, we also investigated whether any of the remaining HEXACO dimensions were linked to cooperation (Table 2) and indeed found some noteworthy evidence for the Emotionality dimension. Specifically, Emotionality was positively linked to cooperation in the chicken game but negatively linked to cooperation in the stag hunt game. This nicely matches the fact that temptation motivates defection in the chicken game (but fear does not) whereas fear motivates defection in the stag hunt game (but temptation does not). In other words, Emotionality is, by definition, closely tied to anxiety and fear (Lee & Ashton, 2004, 2006), which, in turn, is reflected in its association

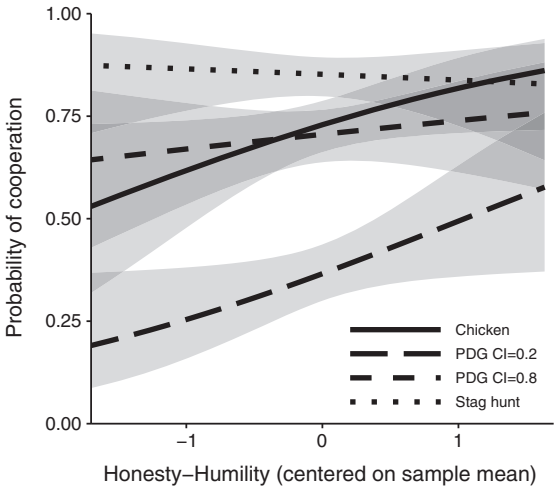

Figure 3. Probability of cooperation depending on Honesty–Humility (centred on the sample mean) and the type of game (Experiment). Confidence bands indicate the 95% confidence intervals. PDG, prisoner's dilemma game

DOI: 10.1002/per

with risky versus non-risky cooperation behaviour. However, it must be noted that this pattern was not statistically robust (although descriptively showing the same trends) in the reanalysis of Kieslich and Hilbig's (2014) experiment as reported earlier (Table 1). Thus, the findings should be interpreted with caution.

GENERAL DISCUSSION

Behaviour in social dilemmas has long been known to yield substantial individual variation (Balliet et al., 2009; Sally, 1995), and a growing body of literature has identified personality traits that account for said variation (Zhao & Smillie, 2015). Among broad basic traits, the Honesty–Humility dimension from the HEXACO model of personality (Ashton et al., 2014; Ashton & Lee, 2007) has repeatedly and most consistently been linked to cooperative behaviour (Hilbig et al., 2012; Mischkowski & Glöckner, 2016; Ruch et al., 2017; Zettler et al., 2013; Zhao & Smillie, 2015). However, the motivational mechanisms underlying this association have largely remained unexplored, because social dilemmas—such as the PDG or public goods game—typically confound temptation to defect (approach) and fear of exploitation (avoidance) as potential determinants of defection. In other words, both greed and fear may underlie a low willingness to cooperate in social dilemma situations (Bruins et al., 1989; Poppe & Utens, 1986; Yamagishi & Sato, 1986), and it is thus unclear which motivation can actually account for said link between Honesty–Humility and cooperation.

In a reanalysis and a new experiment, we set out to separate these motivational mechanisms by manipulating the pay-off structure in a two-person, symmetric, simultaneous game. Specifically, we contrasted pay-off matrices corresponding to the classical PDG structure (in which both temptation and fear imply defection) with matrices representing the chicken game (in which only temptation implies defection) and the stag hunt game (in which only fear implies defection), respectively. Results consistently revealed that Honesty–Humility predicts cooperation versus defection in the PDG if and only if the game is cooperation-unfriendly (i.e. yields a high level of temptation and low level of risk associated with defection). Concurrently, Honesty–Humility predicts cooperation versus defection in the chicken game, whereas it does not explain variance in the stag hunt game or a cooperation-friendly PDG (with a low level of temptation and high level of risk associated with defection). In other words, Honesty–Humility is exclusively linked to cooperation in those games in which temptation may drive defection, but not in games in which only fear would do so. Consequently, results confirm that individuals low in Honesty–Humility are particularly susceptible to the temptation of exploiting others (but not particularly sensitive for the risk inherent in cooperation), whereas their counterparts high in Honesty–Humility are less responsive to the tempting character of defection. Arguably, this is well aligned with the definition of Honesty–Humility and also consistent with other findings showing that low levels of Honesty–Humility entail adapting one's behaviour to situational circumstances in order to maximize one's individual utility (Zettler & Hilbig, 2010; Zettler & Hilbig, 2015).

On the level of more general implications, our findings are aligned with the argument that Honesty–Humility subsumes the variance between people in *seeking* to exploit others. In line with recent ideas on situational affordances —different opportunities to express behaviour, and thereby aspects of personality (Rauthmann, 2012)—Honesty–Humility indeed seems inherently tied to situations in which one may or may not be tempted to maximize one's gains at the cost of others. Within the DIAMONDS taxonomy of situational characteristics (Rauthmann et al., 2014), for instance, Honesty–Humility in particular is linked to situations that provide the opportunity to deceive someone (Sherman, Rauthmann, Brown, Serfass, & Jones, 2015). In a similar vein, De Vries, Tybur, Pollet, and van Vugt (2016) directly map Honesty–Humility onto the domain-specific situational affordance of exploitation. Importantly, in combination with our findings, these links also imply that individuals low in Honesty–Humility should be more likely to *seek* situations that offer an opportunity for exploitation. Thus, specific hypotheses can be derived and tested in future experiments in which one does not merely observe the choices of individuals placed *in* a specific situation (as practically all research on Honesty–Humility has done, to the best of our knowledge), but rather their choices *of* certain situations.

The link between low levels in Honesty–Humility and greedy, selfish or, more generally, antagonistic behaviour is also in line with this dimensions' negative relations with dark traits such as Machiavellianism, Narcissism and Psychopathy (Lee & Ashton, 2005; Muris, Merckelbach, Otgaar, & Meijer, 2017) and indeed the general dispositional tendency of which many different dark traits are specific manifestations, the 'dark core of personality' (Moshagen, Hilbig, & zettler, in press). Our findings are consistent with studies investigating the effects of dark traits in decision-making processes. Specifically, Berg, Lilienfeld, and Waldman (2013) investigated the role of different dark traits on several economic tasks, including a PDG. They found that fearless dominance, a facet of Psychopathy (Lilienfeld & Andrews, 1996), was not related to any criterion (except for one ultimatum game variant). In contrast, both Machiavellianism and overall Psychopathy, which can be considered proxies of low Honesty–Humility, negatively predicted cooperation in the PDG.

On a methodological level, the present work further demonstrates the usefulness of economic game paradigms not only to study actual cooperative behaviour but also to gain a more fine-grained understanding of what drives said behaviour (Thielmann et al., 2015). Given that economic games consist of well-defined pay-off structures that can be analysed and described in terms of Game Theory (Luce & Raiffa, 1957; von Neumann & Morgenstern, 1944) and Interdependence Theory (Kelley & Thibaut, 1978; Thibaut & Kelley, 1959), they provide a valuable toolbox to understand overt behaviour as well as the underlying motivational mechanisms and individual differences therein. Economic games will be most informative whenever different variants are combined or, within a game, pay-off structures are

DOI: 10.1002/per

experimentally manipulated so as to isolate certain behavioural tendencies, which can then be linked to personality (or other variables) in a person–situation–interaction framework.

However, it must be acknowledged that experiments linking (self-report) measures of personality to behaviour in economic games tend to involve relatively small effect sizes —especially if interactions are of prime interest—and thus even relatively large sample sizes provide only limited statistical power. Indeed, there are several reasons why large effects typically cannot be expected in such setups: Games, unlike personality scales, are highly situation-specific and often played with only a few trials (limiting their reliability). Also, self-reported personality measures and behavioural games will have very little (if any) common method variance. Finally, games—unlike personality questionnaires—counteract socially desirable responding by means of anonymity and choice-contingent incentives (Hilbig et al., 2015; Thielmann, Heck, & Hilbig, 2016). Taken together, small effects are to be expected (Hilbig, Thielmann, Klein, & Henninger, 2016). Indeed, several of the effects we report can be considered small at best, and thus more research with substantially larger sample sizes and more statistical power would seem highly advisable.

Moreover, specific research questions concerning the motivational drivers of cooperation may benefit from alternative approaches beyond such combinations of games or game variants. For example, additional insight could be gained from asking participants to predict the other player's choice, which could be taken as an indication of whether a specific motivation may be at play (e.g. if one expects the other to cooperate in the PDG, defection cannot be due to fear). Extending this type of approach, Butler et al. (2011) used thorough interviews and directly asked participants about their motivations in the game, their expectations of the other player, and their perception of the game. Similar but more implicit approaches might be to assess individuals' mental representation of social dilemma situations and similar games by asking them what they consider the best and worst outcome in a game (Halevy, Chou, & Murnighan, 2012; Halevy, Cohen, Chou, Katz, & Panter, 2014) or by assessing the perceived interdependence between players (Gerpott, Balliet, Columbus, Molho, & De Vries, in press).

Alternative approaches such as the ones sketched earlier should arguably be further complemented by future investigations extending the present research in terms of different samples (especially including non-'weird' populations, see Henrich, Heine, & Norenzayan, 2010), alternative measures of personality and traits related to Honesty–Humility (especially Social Value Orientations, e.g. Ackermann et al., 2016; Murphy & Ackermann, 2014), and alternative real-life cooperation situations. As it stands, the present research is limited to a typical Western, educated student sample, self-reported personality, and the arguable artificiality of lab-based economic games. Therefore, the present findings should not be prematurely generalized but instead be taken as a first experimental test of potential motivational mechanisms underlying the overt link between Honesty–Humility and cooperation.

In conclusion, the present work confirms not only earlier observations that individual differences in Honesty–Humility account for cooperation versus defection in social dilemmas and economic decisions more generally but also—and more importantly—*when* and thereby *why* this will be the case: This broad, basic trait predicts (non)cooperative behaviour specifically if the situation entails a strong temptation to exploit others, corroborating the view that those low in Honesty–Humility are motivated by and sensitive to the tempting character of decision-making situations. In turn, once temptation is removed from (or substantially reduced in) the situation, they tend to cooperate to an extent that is comparable with that in individuals high in Honesty–Humility —even if cooperation is risky, such that being fearful (of exploitation) would imply defection.

ACKNOWLEDGEMENT

The work reported herein was supported by a grant to the first author from the Baden-Württemberg Foundation.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

REFERENCES

Ackermann, K. A., Fleiß, J., & Murphy, R. O. (2016). Reciprocity as an individual difference. *Journal of Conflict Resolution*, *60*, 340–367. https://doi.org/10.1177/0022002714541854.

Ahn, T.-K., Ostrom, E., Schmidt, D., Shupp, R., & Walker, J. (2001). Cooperation in PD games: Fear, greed, and history of play. *Public Choice*, *106*, 137–155. https://doi.org/10.1023/A:1005219123532.

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, *11*, 150–166. https://doi.org/10.1177/1088868306294907.

Ashton, M. C., & Lee, K. (2008). The HEXACO model of personality structure and the importance of the H factor. *Social and Personality Psychology Compass*, *2*, 1952–1962. https://doi.org/10.1111/j.1751-9004.2008.00134.x.

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, *91*, 340–345. https://doi.org/10.1080/00223890902935878.

Ashton, M. C., Lee, K., & De Vries, R. E. (2014). The HEXACO Honesty–Humility, Agreeableness, and Emotionality factors: A review of research and theory. *Personality and Social Psychology Review*, *18*, 139–152. https://doi.org/10.1177/1088868314523838.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64*, 359–372. https://doi.org/10.1037/h0043445.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*, 1390–1396. https://doi.org/10.1126/science.7466396.

Balliet, D., Parks, C., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes Intergroup Relations*, *12*, 533–547. https://doi.org/10.1177/1368430209105040.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on*

DOI: 10.1002/per

Psychological Science, *2*, 396–403. https://doi.org/10.1111/j.1745-6916.2007.00051.x.

Berg, J. M., Lilienfeld, S. O., & Waldman, I. D. (2013). Bargaining with the devil: Using economic decision-making tasks to examine the heterogeneity of psychopathic traits. *Journal of Research in Personality*, *47*, 472–482. https://doi.org/10.1016/j.jrp.2013.04.003.

Bruins, J. J., Liebrand, W. B., & Wilke, H. A. (1989). About the saliency of fear and greed in social dilemmas. *European Journal of Social Psychology*, *19*, 155–161. https://doi.org/10.1002/ejsp.2420190207.

Butler, D. J., Burbank, V. K., & Chisholm, J. S. (2011). The frames behind the games: Player's perceptions of prisoners dilemma, chicken, dictator, and ultimatum games. *The Journal of Socio-Economics*, *40*, 103–114. https://doi.org/10.1016/j.socec.2010.12.009.

Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral & Brain Sciences*, *26*, 139–198.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, *31*, 169–193. https://doi.org/10.1146/annurev.ps.31.020180.001125.

de Heus, P., Hoogervorst, N., & Dijk, E. v. (2010). Framing prisoners and chickens: Valence effects in the prisoner's dilemma and the chicken game. *Journal of Experimental Social Psychology*, *46*, 736–742. https://doi.org/10.1016/j.jesp.2010.04.013.

De Vries, R. E., Tybur, J. M., Pollet, T. V., & van Vugt, M. (2016). Evolution, situational affordances, and the HEXACO model of personality. *Evolution and Human Behavior*, *37*, 407–421. https://doi.org/10.1016/j.evolhumbehav.2016.04.001.

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, *34*, 169–189. https://doi.org/10.1207/s15326985ep3403_3.

Elliot, A. J., & Covington, M. V. (2001). Approach and avoidance motivation. *Educational Psychology Review*, *13*, 73–92. https://doi.org/10.1023/A:1009009018235.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*, 171–178. https://doi.org/10.1007/s10683-006-9159-4.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, *6*, 347–369. https://doi.org/10.1006/game.1994.1021.

Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software*, *8*(15), 1–27.

Funder, D. C. (2001). Personality. *Annual Review of Psychology*, *52*, 197–221. https://doi.org/10.1146/annurev.psych.52.1.197.

Funder, D. C. (2009). Naive and obvious questions. *Perspectives on Psychological Science*, *4*, 340–344. https://doi.org/10.1111/j.1745-6924.2009.01135.x.

Gächter, S. (2004). Behavioral game theory. In D.J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 485–503). Malden, MA: Blackwell Publishing, DOI: https://doi.org/10.1002/9780470752937.ch24.

Gerpott, F. H., Balliet, D., Columbus, S., Molho, C., & de Vries, R. E. (2017) (in press). How do people think about interdependence? A multidimensional model of subjective outcome interdependence. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000166.

Glöckner, A., & Hilbig, B. E. (2012). Risk is relative: Risk aversion yields cooperation rather than defection in cooperation-friendly environments. *Psychonomic Bulletin & Review*, *19*, 546–553. https://doi.org/10.3758/s13423-012-0224-z.

Haesevoets, T., Folmer, C. R., & Van Hiel, A. (2015). Cooperation in mixed-motive games: The role of individual differences in selfish and social orientation. *European Journal of Personality*, *29*, 445–458. https://doi.org/10.1002/per.1992.

Halevy, N., Chou, E. Y., & Murnighan, J. K. (2012). Mind games: The mental representation of conflict. *Journal of Personality and Social Psychology*, *102*, 132–148. https://doi.org/10.1037/a0025389.

Halevy, N., Cohen, T. R., Chou, E. Y., Katz, J. J., & Panter, A. T. (2014). Mental models at work: Cognitive causes and consequences of conflict in organizations. *Personality and Social Psychology Bulletin*, *40*, 92–110. https://doi.org/10.1177/0146167213506468.

Henninger, F., Kieslich, P. J., & Hilbig, B. E. (2017). Psynteract: A flexible, cross-platform, open framework for interactive experiments. *Behavior Research Methods*, *49*, 1605–1614. https://doi.org/10.3758/s13428-016-0801-6.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61–83. https://doi.org/10.1017/S0140525X0999152X.

Hilbig, B. E., Glöckner, A., & Zettler, I. (2014). Personality and pro-social behavior: Linking basic traits and social value orientations. *Journal of Personality and Social Psychology*, *107*, 529–539. https://doi.org/10.1037/a0036074.

Hilbig, B. E., Thielmann, I., Hepp, J., Klein, S. A., & Zettler, I. (2015). From personality to altruistic behavior (and back): Evidence from a double-blind dictator game. *Journal of Research in Personality*, *55*, 46–50. https://doi.org/10.1016/j.jrp.2014.12.004.

Hilbig, B. E., Thielmann, I., Klein, S. A., & Henninger, F. (2016). The two faces of cooperation: On the unique role of HEXACO Agreeableness for forgiveness versus retaliation. *Journal of Research in Personality*, *64*, 69–78. https://doi.org/10.1016/j.jrp.2016.08.004.

Hilbig, B. E., Zettler, I., & Heydasch, T. (2012). Personality, punishment, and public-goods: Strategic shifts towards cooperation as a matter of dispositional Honesty–Humility. *European Journal of Personality*, *26*, 245–254. https://doi.org/10.1002/per.830.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence*. New York, NY, US: Wiley.

Kieslich, P. J., & Henninger, F. (2017). Mousetrap: An integrated, open-source mouse-tracking package. *Behavior Research Methods*, *49*, 1652–1667. https://doi.org/10.3758/s13428-017-0900-z.

Kieslich, P. J., & Hilbig, B. E. (2014). Cognitive conflict in social dilemmas: An analysis of response dynamics. *Judgment and Decision making*, *9*, 510–522.

Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology*, *24*, 183–214. https://doi.org/10.1146/annurev.soc.24.1.183.

Krueger, J. I. (2007). From social projection to social behaviour. *European Review of Social Psychology*, *18*, 1–35. https://doi.org/10.1080/10463280701284645.

Krueger, J. I., & Acevedo, M. (2007). Perceptions of self and other in the prisoner's dilemma: Outcome bias and evidential reasoning. *American Journal of Psychology*, *120*, 593–618.

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research*, *39*, 329–358. https://doi.org/10.1207/s15327906mbr3902_8.

Lee, K., & Ashton, M. C. (2005). Psychopathy, machiavellianism, and narcissism in the five-factor model and the HEXACO model of personality structure. *Personality and Individual Differences*, *38*, 1571–1582. https://doi.org/10.1016/j.paid.2004.09.016.

Lee, K., & Ashton, M. C. (2006). Further assessment of the HEXACO Personality Inventory: Two new facet scales and an observer report form. *Psychological Assessment*, *18*, 182–191. https://doi.org/10.1037/1040-3590.18.2.182.

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal population. *Journal of Personality Assessment*, *66*, 488–524. https://doi.org/10.1207/s15327752jpa6603_3.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. New York, NY, US: John Wiley & Sons.

Lüdecke, D. (2017). *sjPlot: Data Visualization for Statistics in Social Science*. https://CRAN.R-project.org/package=sjPlot

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*, 314–324. https://doi.org/10.3758/s13428-011-0168-7.

McClelland, D. C. (1951). *Personality*. New York, NY: William Sloane Assoc, DOI: https://doi.org/10.1037/10790-000.

Mischkowski, D., & Glöckner, A. (2016). Spontaneous cooperation for prosocials, but not for proselfs: Social value orientation moderates spontaneous cooperation behavior. *Scientific Reports*, *6*, 21555. https://doi.org/10.1038/srep21555.

Moshagen, M., Hilbig, B. E., & Musch, J. (2011). Defection in the dark? A randomized-response investigation of cooperativeness in social dilemma games. *European Journal of Social Psychology*, *41*, 638–644. https://doi.org/10.1002/ejsp.793.

Moshagen, M., Hilbig, B. E., & Zettler, I. (in press). The dark core of personality. *Psychological Review*.

Moshagen, M., Hilbig, B. E., & Zettler, I. (2014). Faktorenstruktur, psychometrische Eigenschaften und Messinvarianz der deutschsprachigen Version des 60-Item HEXACO Persönlichkeitsinventars [Factor structure, psychometric properties, and measurement invariance of the German-language version of the 60-item HEXACO personality inventory]. *Diagnostica*, *60*, 86–97.

Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The malevolent side of human nature: A meta-analysis and critical review of the literature on the dark triad (narcissism, Machiavellianism, and psychopathy). *Perspectives on Psychological Science*, *12*, 183–204. https://doi.org/10.1177/1745691616666070.

Murphy, R. O., & Ackermann, K. A. (2014). Social value orientation: Theoretical and measurement issues in the study of social preferences. *Personality and Social Psychology Review*, *18*, 13–41. https://doi.org/10.1177/1088868313501745.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, *314*, 1560–1563. https://doi.org/10.1126/science.1133755.

Poppe, M., & Utens, L. (1986). Effects of greed and fear of being gypped in a social dilemma situation with changing pool size. *Journal of Economic Psychology*, *7*, 61–73. https://doi.org/10.1016/0167-4870(86)90012-7.

R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rapoport, A., & Chammah, A. (1965). *Prisoner's dilemma: A study in conflict and cooperation*. Ann Arbor: University of Michigan Press, DOI: https://doi.org/10.3998/mpub.20269.

Rapoport, A., & Chammah, A. (1969). The game of chicken. In I. R. Buchler, & H. G. Nuttini (Eds.), *Game theory in the behavioral sciences* (pp. 151–175). Pittsburg: University of Pittsburgh Press.

Rapoport, A., & Eshed-Levy, D. (1989). Provision of step-level public goods: Effects of greed and fear of being gypped. *Organizational Behavior and Human Decision Processes*, *44*, 325–344. https://doi.org/10.1016/0749-5978(89)90012-5.

Rauthmann, J. F. (2012). You say the party is dull, I say it is lively: A componential approach to how situations are perceived to disentangle perceiver, situation, and perceiver × situation variance. *Social Psychological and Personality Science*, *3*, 519–528. https://doi.org/10.1177/1948550611427609.

Rauthmann, J. F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C. S., Sherman, R. A., … Funder, D. C. (2014). The Situational Eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology*, *107*, 677–718. https://doi.org/10.1037/a0037250.

Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of Social Service Research*, *21*, 37–59. https://doi.org/10.1300/J079v21n04_02.

Ruch, W., Bruntsch, R., & Wagner, L. (2017). The role of character traits in economic games. *Personality and Individual Differences*, *108*, 186–190. https://doi.org/10.1016/j.paid.2016.12.007.

Rusch, H., & Luetge, C. (2016). Spillovers from coordination to cooperation: Evidence for the interdependence hypothesis? *Evolutionary Behavioral Sciences*, *10*, 284–296. https://doi.org/10.1037/ebs0000066.

Sally, D. (1995). Conversation and cooperation in social dilemmas. A meta-analysis of experiments from 1958 to 1992. *Rationality and Society*, *7*, 58–92. https://doi.org/10.1177/1043463195007001004.

Sherman, R. A., Rauthmann, J. F., Brown, N. A., Serfass, D. G., & Jones, A. B. (2015). The independent effects of personality and situations on real-time expressions of behavior and emotion. *Journal of Personality and Social Psychology*, *109*, 872–888. https://doi.org/10.1037/pspp0000036.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. Piscataway, NJ, US: Transaction Publishers.

Thielmann, I., Böhm, R., & Hilbig, B. E. (2015). Different games for different motives: Comment on Haesevoets, Folmer, and Van Hiel (2015). *European Journal of Personality*, *29*, 506–508. https://doi.org/10.1002/per.2007.

Thielmann, I., Heck, D. W., & Hilbig, B. E. (2016). Anonymity and incentives: An investigation of techniques to reduce socially desirable responding in the trust game. *Judgment and Decision making*, *11*, 527.

Thielmann, I., & Hilbig, B. E. (2014). Trust in me, trust in you: A social projection account of the link between personality, cooperativeness, and trustworthiness expectations. *Journal of Research in Personality*, *50*, 61–65. https://doi.org/10.1016/j.jrp.2014.03.006.

Van Lange, P. A. M. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology*, *77*, 337–349. https://doi.org/10.1037/0022-3514.77.2.337.

Van Lange, P. A. M., De Bruin, E. M. N., Otten, W., & Joireman, J. A. (1997). Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *Journal of Personality and Social Psychology*, *73*, 733–746. https://doi.org/10.1037/0022-3514.73.4.733.

Vlaev, I., & Chater, N. (2006). Game relativity: How context influences strategic decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 131–149.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ, US: Princeton University Press.

Yamagishi, T., & Sato, K. (1986). Motivational bases of the public goods problem. *Journal of Personality and Social Psychology*, *50*, 67–73. https://doi.org/10.1037/0022-3514.50.1.67.

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics*, *6*, 299–310. https://doi.org/10.1023/A:1026277420119.

Zettler, I., & Hilbig, B. E. (2010). Honesty–Humility and a person–situation–interaction at work. *European Journal of Personality*, *24*, 569–582.

Zettler, I., & Hilbig, B. E. (2015). Honesty–Humility. In J. D. Wright (Eds.), *International encyclopedia of social and behavioral sciences* (2nd ed.) Vol. 11, pp. 169–174). Oxford, UK: Elsevier, https://doi.org/10.1016/B978-0-08-097086-8.25068-X.

Zettler, I., Hilbig, B. E., & Heydasch, T. (2013). Two sides of one coin: Honesty–Humility and situational factors mutually shape social dilemma decision making. *Journal of Research in Personality*, *47*, 286–295. https://doi.org/10.1016/j.jrp.2013.01.012.

Zhao, K., Ferguson, E., & Smillie, L. D. (2017). Individual differences in good manners rather than compassion predict fair allocations of wealth in the dictator game. *Journal of Personality*, *85*, 244–256. https://doi.org/10.1111/jopy.12237.

Zhao, K., & Smillie, L. D. (2015). The role of interpersonal traits in social decision making exploring sources of behavioral heterogeneity in economic games. *Personality and Social Psychology Review*, *19*, 277–302. https://doi.org/10.1177/1088868314553709.