MDPI

Article

Dark Triad, Impulsiveness and Honesty-Humility in the Prisoner's Dilemma Game: The Moderating Role of Gender

Olga Lainidi [1,*] , Eirini Karakasidou [2] and Anthony Montgomery [3]

1 School of Psychology, University of Central Lancashire, Fylde Rd, Preston PR1 2HE, UK
2 School of Psychology, Panteion University, 176 71 Athens, Greece
3 School of Psychology, Northumbria University, Newcastle upon Tyne NE1 8ST, UK
* Correspondence: lndolga@gmail.com

Abstract: (1) Background: The aim of the study was to investigate if the Dark Triad (DT)—which includes psychopathy, Machiavellianism and narcissism—impulsiveness and Honesty-Humility (HH), can predict individuals' intended behavior in a one-shot Prisoner's Dilemma Game (PDG) and whether this relationship is moderated by gender. (2) Methods: A cross-sectional correlational design was used, examining regression and moderation models. A total sample of 197 working adults from Greece (64% women, M_{age} = 35.13 years old) completed a one-shot, simulated PDG, the Dirty Dozen scale, the Barratt Impulsiveness Scale and the Honesty-Humility subscale from the HEXACO NEO-PI. (3) Results: Significant correlations between overall DT score, narcissism and psychopathy scores and impulsiveness and all three DT traits, as well as the overall DT score and honesty-humility, were identified. Overall DT scores were found to increase the odds of defecting, while gender significantly moderated the effects of Machiavellianism, HH and impulsiveness on the participants' behavior in the one shot PDG, leading to different effects for men and women. (4) Conclusions: Gender moderates the relationship between certain personality traits and behavior in social situations involving interdependence and/or moral decision making, such as the PDG, leading to changes both in terms of the statistical significance and the direction of the effect. Simulated social situations could serve as situational judgment tests in an effort to develop a better understanding of the underpinning mechanisms between personality, gender and social behavior.

Keywords: Dark Triad; impulsiveness; Honesty-Humility; the Prisoner's Dilemma Game

1. Introduction

Mathematical game theory offers models for studying ethical decision making and intended moral behavior in in-vitro social situations [1,2]. In such a framework, one individual is asked to choose among a range of options based on the known or expected behavior of another individual or other individuals. The outcomes and the consequences of the choices depend on the decisions of the other individual(s) and are represented with prefixed and clear numerical values. Competition and cooperation are two concepts that occupy a large place in the field of human sciences. The "prisoner's dilemma" is a well-known game, a cooperation problem, where the outcome is determined by the players' decisions to either cooperate with their partner or betray them [3]. It is a two-person, non-zero-sum game where a graded pattern of rewards is offered based on the constellation of decisions made by the two players [4,5]. In the Prisoner's Dilemma Game (PDG), there is a temptation for the player to trick the other player, as a successful defection leads to the highest reward possible, whilst there is a smaller reward for cooperation, as there would be no reason for defection if the reward for mutual cooperation was higher. Meanwhile, the lowest reward is offered when the individual is being defected by the opponent [6,7]. The interest in the relationship between personality, moral judgment and intended behavior in ethical dilemmas is deeply rooted in the history of philosophy and psychology. In this study, we examined

the potential role of two normal-range personality traits—meaning Honesty-Humility (HH) and impulsiveness—and the Dark Triad personality traits—Machiavellianism, narcissism and psychopathy—in an online simulation of a one-shot PDG.

The Dark Triad (DT) is a structure of three theoretically distinct (though empirically overlapping) personality characteristics (psychopathy, Machiavellianism and narcissism) that are largely characterized by interpersonal maladaptation [8,9] and fall under the umbrella of antagonism [10]. Psychopathy, as part of the DT, is characterized by a lack of conscience and guilt [11,12], as well as uncooperative behavior in a variety of social situations [13,14], and it has been suggested that lower levels of cooperativeness pay higher rewards for individuals scoring higher in psychopathy [15]. The second traits of the DT, Machiavellianism, is a psychological notion that was developed approximately during 1954–1955 and has been identified as the core of manipulation with four main characteristics: lack of emotional affection within interpersonal relationships, lack of interest in conventional morality, lack of psychopathology awareness and lack of ideological commitment [16]. The third trait of the DT is narcissism, a term first used by Havelock Ellis [17] to describe a clinical condition of auto-eroticism. Narcissism is a relatively constant differentiation of human personality characterized mainly by a sense of magnificence, overrated sensation of beauty and falsified views of oneself, extending to individual, interpersonal relationships and self-regulation strategies, as well as self-enhancement and social dominance [18,19]. The DT exhibits strong correlations with self-reported measures of impulsiveness [20,21], a trait that has also been related to the likelihood of defecting in the PDG [22]. Generally, impulsiveness is a concept related to externalized behaviors that seem to have not been well thought through [23]. Barratt [24] conceptualized impulsiveness into three components: motor impulsivity, cognitive impulsivity and the difficulty of planning over time or anticipation.

Honesty-Humility (HH) from the HEXACO model of personality [25] has been referred to as a core opposite trait of the DT, with significant, high correlations between the two [10]. HH has been found to be associated with adjectives such as fair, generous and modest while being opposed to adjectives such as dishonest, unfair, greedy and boastful [26] and is considered to be associated with altruism and/or cooperation via a reluctance to exploit others even when one could get away with doing it [27]. Differences in HH are associated with individual differences in the tendency to be authentic at the interpersonal level, to avoid deception and corruption, to not care about social status and wealth, to be humble and modest and to be willing to give up personal gain for collective benefit [28,29].

The potential role of the aforementioned variables in the PDG has been examined in both one-shot and iterated simulations. Deutchman and Sullivan [30] found that participants with high DT scores were more likely to behave selfishly in a one-shot PDG. Congruently, Malesza [22] reported that the DT could significantly predict the likelihood of defection in an iterated PDG with 10 consecutive rounds; more specifically, psychopathy and Machiavellianism but not narcissism were identified as significant predictors of the participants' likelihood of defecting, while high impulsiveness combined with high psychopathy could better predict defection. Conversely, in another study, it was reported that narcissism was related to initial cooperative behavior eventually followed by increased defective behavior in iterated PDGs [31]. Curry et al. [32], in their study, showed that individuals high in Machiavellianism were less likely to cooperate as a response or to initiate cooperation. On the other hand, HH has been suggested as a relatively stable predictor of cooperative behavior. In particular, Hilbig et al. [33], in their study, found that HH was a significant predictor of cooperation in cooperation games, especially in cases where temptation was a possible motivational factor for defection. These findings suggest that individuals with high levels of HH have more potential to resist a temptation that motivates them to defect compared to those with lower levels of HH. Zettler et al. [34] also explored the role of HH in social decision making and found that the changes in participants' decisions greatly depended on the dispositional level of HH when punishment was introduced. HH was linked to prosocial behavior in the dictator giving game [35] and the chicken

game [33], although there have also been reports of inconsistency in the relationship between HH and prosocial behaviors in social situations that allow exploitation of others [36]. The aforementioned literature and empirical evidence informed Hypotheses 1 and 2 of the study, which were formulated as follows:

H1a. *There is a statistically significant negative correlation between HH and the DT.*

H1b. *There is a statistically significant positive correlation between impulsiveness and the DT.*

H2. *The DT, HH and impulsiveness are significant predictors of the participants' behavior in the PDG.*

Gender differences in personality are well documented in the empirical literature [37–39] and have also been reported with regards to the DT personality traits [40] and the motivational behavior control aspect of impulsiveness [41]. However, the literature on behavior in cooperation games such as the PDG is less clear in terms of gender differences; while women are often stereotyped to be more cooperative and communal than men [42], findings from PDG studies have shown that women are often less cooperative in fully incentivized one-shot PDGs [43] as well as when risk is high [44]. This is in agreement with empirical evidence that suggests that women are more risk averse than men in social risk taking [45], in objective probability gambles [46] and in financial risks [47]. In addition, as the PDG involves fear of exploitation and expectations regarding other peoples' cooperation [43], it is still not clear whether the stereotypically higher prosociality of women is translated into cooperative behavior in social dilemmas and whether the same personality traits have similar effects on men's and women's behavior in terms of magnitude and direction. Sociocultural theory [48] has been previously used to explain gender differences in personality and emotion, postulating that gender differences arise from social, cultural, psychological and other environmental forces while acknowledging the roles of biological effects and learned influences [49]. Thus, we expected gender to moderate the relationship between the examined personality traits and the likelihood of the participants defecting in the one-shot PDG (Figure 1). The aforementioned literature and empirical evidence informed Hypothesis 3 of the study, which was formulated as follows:
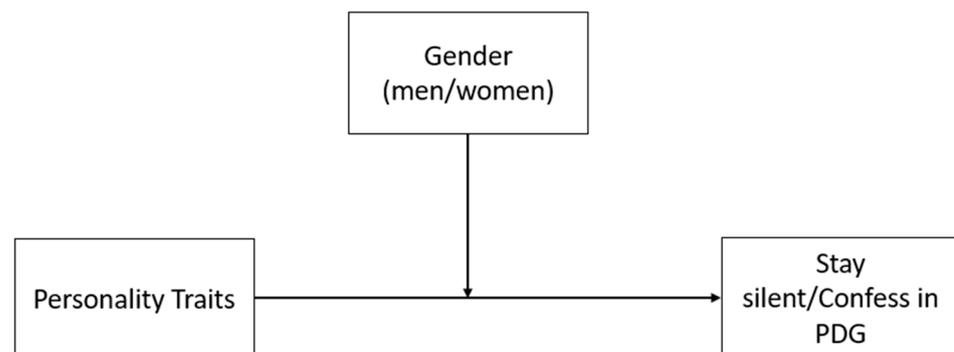
Figure 1. Hypothesized model for the moderating effect of gender on the relationship between the examined personality traits and participants' behavior in the PDG.

H3. *Gender moderates the relationship between (a) the DT, (b) HH, (c) impulsiveness and the participants' behavior in the PDG.*

2. Materials and Methods

2.1. Design

The research followed a quantitative, cross-sectional correlational design. In particular, participants' behavior in the PDG was the dependent variable in the predictive models with two levels, defect (confess) and cooperate (remain silent). The DT scores, impulsiveness and HH were the independent variables/predictors.

2.2. Participants

A total sample of 197 individuals (N = 197, 64% women, M_{age} = 35.13 years old) participated in the study. Power analysis indicated a minimum of 85 participants in order to detect significant effects with a statistical power of 0.80 (Pr(Y = 1/X = 1)H1 = 0.60). Opportunity/Convenience sampling was used to collect responses from a general population sample (Table 1, demographic information). Inclusion criteria were: participants had to be adults (>18 years old) and fluent in the Greek language (language of the survey). Exclusion criteria were: individuals with visual disabilities were excluded as the survey was only available in written form, and individuals with diagnosed mental health disorders related to personality were excluded as well due to the use of personality trait measures, which potentially refer to aspects that might overlap with symptoms of mental health disorders (e.g., the DT and Antisocial Personality Disorder). Further demographic information is presented in Table 1, indicating that the sample was generally heterogeneous, similar to relevant studies that have examined the issue in the general population.

Table 1. Demographic information.

		F	**%**
Education	Junior high school	7	3.6
	High school	22	11.2
	Technical education	42	21.3
	University/college	65	33.0
	Master's degree	50	25.4
	Doctorate degree	11	5.6
Marital status	Unmarried	106	53.8
	Married	44	22.3
	Divorced	33	16.8
	Cohabitation agreement	4	2.0
	Other	10	5.1
Employment status	Employee in private sector	101	51.3
	Employee in public sector	24	12.1
	Freelancer	50	25.4
	Other	22	11.2

2.3. Measures

The participants completed an online questionnaire that consisted of a one-shot PDG, three standardized scales and one improvised demographics questionnaire. The measures administered to the participants were the Dirty Dozen scale for measuring the DT, the Barratt Impulsiveness Scale and the Honesty-Humility subscale from the HEXACO NEO-PI personality inventory.

Adaptation of the PDG: The PDG was adapted as a one-shot item for the purposes of the study. In particular, the participants were informed about a fictional situation in which they were required to imagine themselves prior to choosing their desired response. They were presented with a short text explaining the scenario of the PDG, and they were invited to play in real time. The participants were informed that each one of them is one of two criminals who were arrested due to allegedly committing a burglary, and they were given the following options: defect (confess) or cooperate (remain silent) (Table 2). In the light of the above, they were asked to choose their response in a one-shot game, and, after the participant chose their response, they were informed that the other participant (predefined option) in fact made the same decision as them.

Table 2. Cross-tabulation of the outcome for each player in the PDG.

Player 1	Player 2	
	Defect	Cooperate
Defect (Confess)	5 years in prison for each	Pl.1: released/Pl.2: 20 years in prison
Cooperate (Remain Silent)	Pl.1: 20 years in prison/Pl.2: released	1 year in prison for each

Dirty Dozen: For the measurement of the DT, the Dirty Dozen scale was used [50], back-translated in Greek. The Dirty Dozen scale consists of 12 items measured on Likert scales, divided into three quadrants, each referring to a different aspect of the dark personality—psychopathy, Machiavellianism and narcissism. Despite its division into three dimensions, the Dirty Dozen scale remains a single scale, yielding a total score as well [50], with a good fit for the bifactorial model [21,50,51]. The Dirty Dozen scale shows high internal consistency levels for the total scale (alpha = 0.83).

Barratt's Impulsiveness Scale: Barratt's Impulsiveness Scale (BIS) [24] is a self-report scale with 30 items. The items are scored on a 4-point scale (1 = rarely/never, 2 = occasionally, 3 = often, 4 = almost always/always). Reliability coefficients for the total score (Cronbach's alpha) range from 0.72 to 0.83. A global score, sum of all the items, is attributed. Each of the subscales (Barratt factors) receives a score according to the items that compose it: motor impulsiveness, measuring the behavioral dimension of impulsiveness; cognitive impulsiveness, related to the potential difficulties experienced in intellectual tasks and their own ability to evaluate their mode of cognitive functioning; and planning difficulty, related to a more general dimension of impulsivity and, in particular, to the overall mode of operation of the subject in relation to the notion of the future and self-control [24].

Honesty-Humility: HH was measured with the relevant subscale from the HEXACO Personality Inventory [25]. The subscale consists of 10 items answered on 5-point Likert scales (1 = totally disagree, 5 = totally agree). The subscale has high internal consistency (alpha = 0.72).

2.4. Procedure

The study was approved by the Ethics Committee of the University of Central Lancashire. Surveys were administered online via the Google Forms platform and shared via social media platforms. Eligible participants received the link so that they could respond online. The participants were informed about the purpose and the context of the study so that they could give their informed consent in order to be able to proceed to further completing the questionnaire. After the consent form, the participants were presented with the PDG scenario in detail and the outcome for each of their choices. The participants completed the one-shot PDG followed by the Dirty Dozen scale, which was followed by the Barratt's Impulsiveness Scale and, afterwards, the Honesty-Humility subscale. The improvised demographics questionnaire was the last part of the form. Participants also were offered the option to submit their email address in case they wanted to be informed about the summative results of the study. Participation was anonymous; data were confidential, and all ethical requirements were met.

2.5. Analysis

Statistical analyses were conducted with SPSS v28. Variables were screened for normal distribution with standardized values of skewness and kurtosis [52], and reliability coefficients were calculated for scales and subscales (Cronbach's alpha). The global score was used for impulsiveness. For the DT, the bifactorial model was taken into account, leading to one overall DT variable as well as one separate variable for each of the three constructs. Participants responses in the PDG simulation were coded as a binary variable with the higher value attributed to defecting (0 = cooperate/remain silent, 1 = defect/confess). Pearson r correlations were used to test the first hypothesis regarding the significant correlations between the DT, HH and impulsiveness. A multiple linear regression was performed to examine HH and impulsiveness as predictors of the DT. Binary logistic regression was used

to test whether the DT, HH and impulsiveness could predict the participants' response in the one-shot PDG. Separate regression models were used for variables with strong correlation coefficients ($r > 0.65$) to avoid multicollinearity. Hypothesis 3—meaning the moderating effect of gender—was tested using moderation analysis with logistic regression and a bootstrapping approach to assess the significance of the effects at differing levels of the moderator [53]. Different moderation models were tested for each of the independent variables to examine all potential interactions of gender with the independent variables of the study, while the remaining variables were entered as covariates. In each model, gender was always the moderator (dichotomous variable, 0 = men, 1 = women), and participants' behavior in the one-shot PDG was always the outcome (0 = cooperate/remain silent, 1 = defect/confess). Due to the dichotomous nature of the outcome variable, moderation analysis with logistic regression was deemed to be the most suitable option.

3. Results

Distribution of variables indicated that the data were positively skewed for psychopathy (z-skewness = 3.38/z-kurtosis = 0.77) and Machiavellianism (z-skewness = 4.65/z-kurtosis = 0.86) and negatively skewed for HH (z-skewness = −4.95/z-kurtosis = 2.30) compared to the +/−3.29 threshold recommended by Field [52]. Logarithmic transformation was used for the positively skewed data using the ln(x) function. Normal Score Transformation (NST) was used for the negatively skewed data. Winsorization was used to deal with outliers [52]. Transformations had no impact on the statistical significance of the findings. Table 3 includes the original values for means, SDs and minimum and maximum scores for each variable as well as the standardized values for skewness/kurtosis.

Table 3. Descriptive statistics for the total sample (N = 197) for Machiavellianism, psychopathy, narcissism, impulsiveness and honesty-humility.

	N	Min	Max	M	SD	Z-Skewness **	Z-Kurtosis **
Dark Triad *	197	12.00	75.00	36.74	12.36	2.71	1.73
Machiavellianism	197	4.00	24.00	9.46	4.58	0.30	2.89
Psychopathy	197	4.00	26.00	11.40	5.41	1.34	2.74
Narcissism	197	4.00	28.00	15.87	5.78	0.74	1.42
Impulsiveness	197	36.00	104.00	65.36	11.37	2.19	2.25
Honesty-Humility	197	12.00	49.00	35.51	7.79	0.03	0.51

* Overall mean score of Machiavellianism, psychopathy and narcissism. ** Values for z-skewness and z-kurtosis after applying the ln(x) transformation for psychopathy and Machiavellianism and the NST transformation for honesty/humility.

Correlations between the variables of the study were examined using the Pearson r correlation coefficient to test H1a and H1b (Table 4). A significant positive correlation was identified between the DT and impulsiveness ($r = 0.28$, $p < 0.001$), and a negative strong correlation was found between the DT and HH ($r = -0.62$, $p < 0.001$). Impulsiveness had a significant negative correlation with HH ($r = -0.36$, $p < 0.001$). Significant positive correlations were identified between impulsiveness and narcissism ($r = 0.24$, $p < 0.001$) and psychopathy ($r = 0.23$, $p < 0.01$) but not Machiavellianism. HH was negatively correlated with each of the three dark traits: narcissism ($r = -0.55$, $p < 0.001$); Machiavellianism ($r = -0.28$, $p < 0.001$); and psychopathy ($r = -0.52$, $p < 0.001$). HH and impulsiveness were examined as predictors of DT, controlling for demographic variables, and the regression model showed that HH was a significant predictor of the DT ($\beta = -0.576$, SE = 0.804) ($t = -8.668$, $p < 0.001$, 95% CI (-8.66, -5.49)) while impulsiveness was not.

Table 4. Correlation (Pearson r) coefficients (N = 197, CI 95%).

		1	2	3	4	5
1.	Dark Triad					
2.	Narcissism	0.80 **				
3.	Machiavellianism	0.63 **	0.22 **			
4.	Psychopathy	0.83 **	0.58 **	0.36 **		
5.	Impulsiveness	0.28 **	0.24 **	0.07	0.23 *	
6.	Honesty/Humility	−0.62 **	−0.55 **	−0.28 **	−0.52 **	−0.36 **

* Correlation is significant at the 0.01 level (two-tailed). ** Correlation is significant at the 0.001 level (two-tailed).

Logistic regression was employed as the regression method for Hypotheses 2 and 3. Logistic regression analysis was chosen as the best fit since the outcome variable was dichotomous; moreover, logistic regression is valuable for predicting the likelihood of an event, as it allows the researchers to predict the odds of occurrence between any two classes [54]. Linearity of independent variables and log odds was tested prior to conducting logistic regressions with the use of the Box–Tidwell test and visual check of the scatterplots between each predictor and its logit values, which indicated that the assumption of linearity was satisfied. No significant outliers were identified in the data after the initial winsorization (no data points with absolute standardized residual values > 3), while absence of multicollinearity was implied as VIF < 5.

Regarding Hypothesis 2, the effect of the DT was statistically significant and positive (b = 0.046, SE = 0.02, p = 0.040, OR = 1.05, 95% CI (0.01, 0.09)), slightly increasing the odds of the participants defecting. Additionally, the significant direct effect of impulsiveness was positive (b = 0.058, SE = 0.03, p = 0.027, OR = 1.06, 95% CI (0.01, 0.11)).

The effect of HH was negative, thus, reducing the odds of defecting, but not statistically significant (b = −0.290, SE = 0.30, p = 0.341, OR = 0.74 = 5, 95% CI (−0.89, 0.31)), and the effect of Machiavellianism was also non-significant but positive (b = 0.432, SE = 0.53, p = 0.412, OR = 1.54, 95% CI (−0.60, 1.47)), as was the effect on narcissism (b = 0.075, SE = 0.05, p = 0.129, OR = 1.08, 95% CI (−0.02, 0.17)) and the effect of psychopathy (b = 1.322, SE = 0.68, p = 0.053, OR = 3.75, 95% CI (−0.02, 2.66)). Thus, only the overall DT score and impulsiveness could predict the odds of defection on a statistically significant level, partially supporting the second hypothesis (H2).

As far as Hypothesis 3 is concerned, the results of the moderation analysis indicated that gender had a significant moderating effect on participants' behavior in the PDG, HH, Machiavellianism and impulsiveness, but not for psychopathy, narcissism and overall DT scores.

More specifically, for HH, the results indicated a significant interaction (b = 0.990, SE = 0.36, p = 0.006, OR = 2.69, 95% CI (0.28, 1.70)). In particular, among women, HH had a statistically significant positive effect, increasing the odds of them defecting in the one-shot PDG (b = 0.701, SE = 0.26, OR = 2.02, p = 0.007, 95% CI (0.19, 1.21)). Among men, the effect of HH was negative—thus, decreasing the odds of defecting—but not statistically significant (b = −0.290, SE = 0.30, OR = 0.75, p = 0.34, 95% CI (−0.89, 0.31)). Figure 2 graphs the interaction, showing the change in the expected probability by gender. Overall, the effect of HH on the participants' behaviors changed both in terms of significance and in terms of direction between men and women.
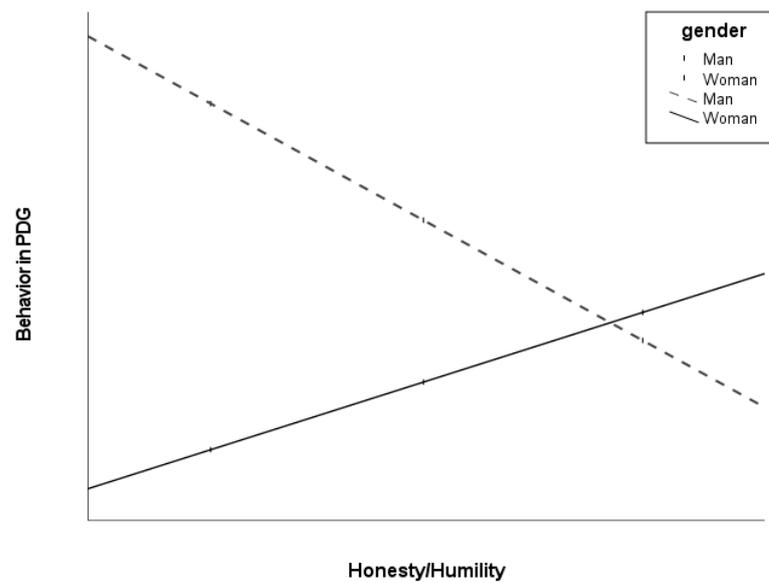
Figure 2. Simple slopes of the regression of honesty/humility on behavior in the PDG at the two levels of gender.
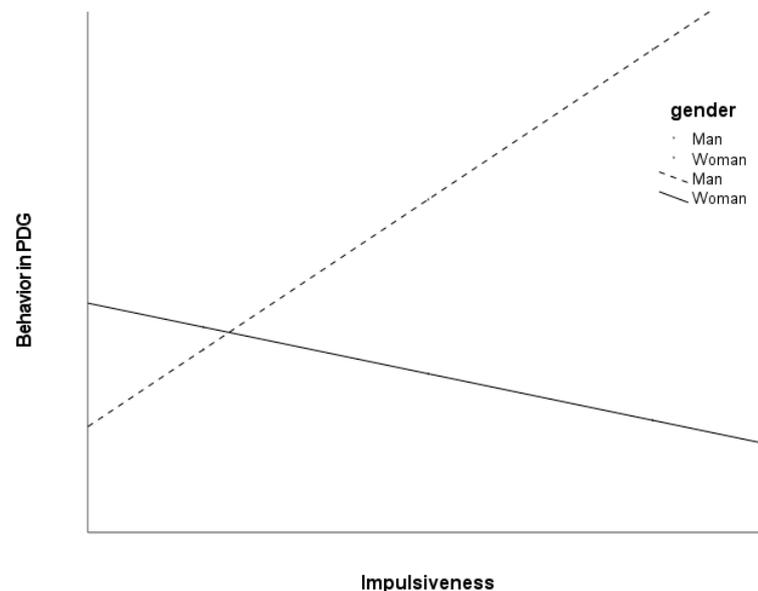
The moderating effect of gender was also significant for the effect of Machiavellianism on participants' behavior in the one-shot PDG, showing a statistically significant interaction, (b = −1.669, SE = 0.69, *p* = 0.015, OR = 0.19, 95% CI (−3.02, −0.32)). In particular, among women, Machiavellianism had a statistically significant negative effect, reducing the odds of defecting in the PDG (b = −1.237, SE = 0.46, OR = 0.29, *p* = 0.007, 95% CI (−2.13, −0.34)). Among men, the effect of Machiavellianism was positive—thus, increasing the odds of defecting—but it was not significant (b = 0.432, SE = 0.53, OR = 1.54, *p* = 0.411, 95% CI (−0.60, 1.47)). Figure 3 graphs the interaction, showing the change in the expected probability by gender. Overall, the effect of Machiavellianism on the participants' behaviors changed both in terms of significance and in terms of direction between men and women.
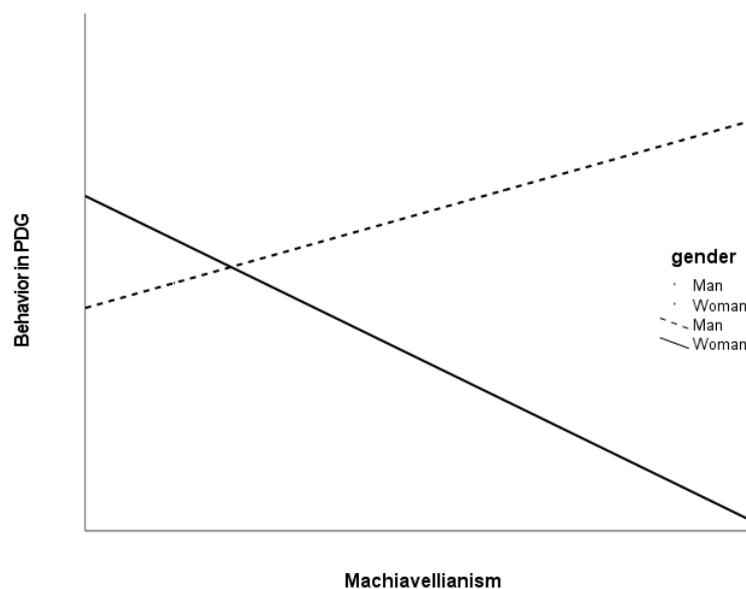
Figure 3. Simple slopes of the regression of Machiavellianism on behavior in the PDG at the two levels of gender.

The moderating effect of gender was also significant for the effect of impulsiveness on participants' behavior in the PDG, showing a statistically significant interaction, (b = −0.078, SE = 0.03, *p* = 0.014, OR = 0.93, 95% CI (−0.14, −0.02)). In particular, among

men, impulsiveness had a statistically significant positive effect, increasing the odds of them defecting in the PDG (b = 0.058, SE = 0.03, OR = 0.29, p = 0.028, 95% CI (0.01, 0.11)). Among women, the effect of impulsiveness was negative but non-significant (b = −0.020, SE = 0.02, OR = 0.98, p = 0.289, 95% CI (−0.06, 0.02)). Figure 4 graphs the interaction, showing the change in the expected probability by gender. Overall, the effect of impulsiveness on the participants' behaviors changed both in terms of significance and in terms of direction between men and women.

Figure 4. Simple slopes of the regression of impulsiveness on behavior in the PDG at the two levels of gender.

The moderating effect of gender was not significant for the effect of the overall DT score on participants' behavior in the PDG (b = −0.036, SE = 0.03, p = 0.193, OR = 0.96, 95% CI (−0.09, −0.02)). No significant moderating effects were identified for the interaction between gender and narcissism (b = −0.009, SE = 0.06, p = 0.877, OR = 0.99, 95% CI (−0.11, 0.10)), while the direct effect of narcissism was also not significant (b = 0.075, SE = 0.05, p = 0.129, OR = 1.08, 95% CI (−0.02, 0.17)). The results of the interaction and the direct effect of psychopathy were non-significant as well (b = −0.962, SE = 0.72, p = 0.182, OR = 0.38, 95% CI (−2.37, 0.45) and b = 1.322, SE = 0.68, p = 0.052, OR = 3.75, 95% CI (−0.02, 2.66), respectively).

4. Discussion

The purpose of this study was to examine the relationship between the DT, impulsiveness, HH and behavior in a one-shot PDG and whether participants' gender was a moderator in these relationships. We hypothesized that HH, impulsiveness and the DT would be significantly correlated (H1ab), that HH, impulsiveness and the DT traits would significantly predict the likelihood of defecting in a one-shot PDG (H2) and that gender would moderate the relationship between the participants' behavior in the PDG and the examined personality traits (H3).

The statistical analyses performed indicated that there was indeed a significant positive correlation between DT and impulsiveness and a significant negative correlation between DT and HH, thus, supporting our first hypothesis (H1ab). Hypothesis 2 (H2) was partially supported, as only the overall DT score and impulsiveness had a significant overall effect on the participants' responses in the PDG. Regarding the third hypothesis (H3), the moderation models tested showed that gender had a significant moderating effect in the cases of Machiavellianism, HH and impulsiveness, but not for the overall DT score, psychopathy or narcissism, thus, partially supporting this hypothesis as well.

The negative correlation between HH and the DT has been well documented in the literature, and it has been suggested that it can be explained via the common low pole that HH shares with the DT [30]; studies have previously explained that there is a part of variance shared between HH and the DT, with strong correlations (-0.80 to -0.94). HH was found to be a significant negative predictor of the DT, in line with previous research which has supported the fact that an increase in HH scores is associated with a decrease in DT scores [30]. Congruently, the positive relationship between the DT traits and impulsiveness is in agreement with previous studies. When examining each of the dark triad traits separately, psychopathy and narcissism were positively correlated to impulsiveness while Machiavellianism was not, in agreement with previous findings where different measures of impulsivity were examined [20,55].

The results of our study also supported the potential effect of the DT and impulsiveness on predicting the likelihood of defection in a one-shotPDG, in agreement with previous research supporting the existing findings that higher DT scores would predict more selfish (defective) behavior [22,30]. However, the moderating effect of gender revealed that, in the cases of HH, Machiavellianism and impulsiveness, the effect of the same traits not only differs in terms of its statistical significance, but also in the direction. More specifically, we found that HH significantly increased the likelihood of confessing among women—i.e., defecting—but the non-significant effect in men had the opposite direction, meaning that it was related to reduced likelihood of defecting. Machiavellianism in women decreased the likelihood of defecting, while, among men, it was found to increase the odds of defection; impulsiveness increased the likelihood of defecting among men but, conversely, reduced the odds of defection among women. Past research has provided support for gender differences in DT traits in terms of how they are expressed; for example, Szabo and Jones [56] found that Machiavellianism was positively associated with planning among men but negatively among women, suggesting that the same traits can be linked to different behavioral patterns. The different effects for men and women that were identified in our study are congruent with the previous literature that identified gender to be a significant moderator in the relationship between personality and behavior-related outcomes, such as counterproductive work behavior [57,58] and buying behavior [59,60], suggesting that personality traits might be operationalized in different ways and, thus, lead to different responses and/or behavioral outcomes among men and women. Such an approach can be supported both from an evolutionary perspective, suggesting that women and men have developed different adaptive mechanisms, and a social-constructionist approach, supporting the fact that these differences are the product of social learning and social influences [43]. As such, findings suggesting that HH leads to an increased likelihood of defecting among women as opposed to decreasing the likelihood of defecting among men could be related to differences in expectations, lower trust and increased risk aversion among women, especially considering that defecting leads to avoiding the worse outcome in the PDG (20 years jailtime if the other person defects) and could also lead to the best outcome if the other person cooperates (0 years jailtime if the other person cooperates).

The findings of the current study are susceptible to certain limitations. In particular, the convenience sampling strategy limits the generalizability of the findings and the representativeness of the studied population, thus, affecting the validity of the research. The present research was also limited by the collection of only self-report data; thus, we do not know the degree to which estimates of the relationships between the variables are biased by self-rating—especially in cases of socially non-desirable traits such as the DT personality traits [61]—and common method variance. Future research should seek to collect data from multiple sources and use different methods of data collection (e.g., experience sampling methodology). In addition, the one-shot PDG gives limited insight into the individuals' behavior patterns compared to an iterated PDG. In addition, we did not examine the effect of contextual factors (e.g., phrasing the game in terms of gain or loss, considering differences between playing against a computer and/or a person, digital simulation versus face-to-face interaction, etc.).

Nonetheless, the study has implications associated with personality assessment and behavior prediction. The implications related to the moderating role of gender support the need to further explore how the same personality traits can potentially lead to different behavioral outcomes and subjective experiences for individuals of different genders. Neglecting the potential differences stemming from demographic variables can result in empirical evidence that either provides more information for particular subgroups or is characterized by very limited external validity and practical usefulness. In the field of management and organizational psychology, these implications extend to increasing cooperation and cohesion in teams, managing counterproductive work behaviors, dealing with unprofessional behaviors, competitiveness and organizational trust and communication, among other things. Moreover, examining the potential effects of personality traits in simulated social situations involving goods, interdependence and/or ethical dilemmas could serve as situational judgment tests in an effort to develop new, more valid and reliable measures of personality and, in particular, to develop a better understanding of the underpinning mechanisms between personality traits and (non)cooperative behavior. Computer-simulated real-life situations associated with personality measures might be an effective and promising method for personality assessment and behavior prediction.

5. Conclusions

There is a need for further investigation of the behavioral phenotypes of the DT personality traits as well examination of possible mediating and moderating effects of key-demographic factors. Moreover, future research should examine the use of virtual reality technologies in decoding real-life social dilemmas, which, combined with personality measures, could allow a more thorough understanding of the complex mechanisms that explain the relations between personality and behavior in social situations involving interdependence and/or ethical decision making. Such studies will allow a deeper insight into the understanding of human personality development and behavior, generating information that can have applications and implications both in theory and in practice, while taking into account key demographic variables that can differentiate the effect that personality traits might have on individual and social behavior.

References

1. Krueger, J.I.; Heck, P.R.; Evans, A.M.; DiDonato, T.E. Social game theory: Preferences, perceptions, and choices. *Eur. Rev. Soc. Psychol.* **2020**, *31*, 222–253. [CrossRef]
2. Luce, R.; Raiffa, H. *Games and Decisions: Introduction and Critical Survey*; Jone Wiley & Sons, Inc.: New York, NY, USA, 1957; Volume 1, p. 958.
3. Axelrod, R. More effective choice in the prisoner's dilemma. *J. Confl. Resolut.* **1980**, *24*, 379–403. [CrossRef]
4. Rasmusen, E.B. Strategic implications of uncertainty over one's own private value in auctions. *Adv. Theor. Econ.* **2006**, *6*, 1–22. [CrossRef]
5. Kim, Y. Application of the Iterated Prisoner's Dilemma and Evolutionary Game Theory in Siting Nuclear Power Plants in South Korea. *Int. J. Math. Models Methods Appl. Sci.* **2022**, *16*, 166–177. [CrossRef]
6. Johnson, D.D.; Stopka, P.; Bell, J. Individual variation evades the Prisoner's Dilemma. *BMC Evol. Biol.* **2002**, *2*, 15. [CrossRef]

7. Mieth, L.; Buchner, A.; Bell, R. Moral labels increase cooperation and costly punishment in a Prisoner's Dilemma game with punishment option. *Sci. Rep.* **2021**, *11*, 10221. [CrossRef]
8. Furnham, A.; Richards, S.C.; Paulhus, D.L. The Dark Triad of Personality: A 10 Year Review. *Soc. Pers. Psychol. Compass* **2013**, *7*, 199–216. [CrossRef]
9. Fehr, B.; Samsom, D. The Construct of Machiavellianism: Twenty Years Later. *Adv. Personal. Assess.* **2013**, *9*, 77.
10. Vize, C.E.; Collison, K.L.; Miller, J.D.; Lynam, D.R. The "core" of the dark triad: A test of competing hypotheses. *Personal. Disord. Theory Res. Treat.* **2020**, *11*, 91. [CrossRef]
11. Hare, R.D.; Neumann, C.S. Psychopathy as a Clinical and Empirical Construct. *Annu. Rev. Clin. Psychol.* **2008**, *4*, 217–246. [CrossRef]
12. Giannini, F.; Raimondi, R. Psychopathological excursus on anti-social personality disorder, psychopathy and the dark triad: A review of international literature. *J. Psychopathol.* **2020**, *26*, 242–247. [CrossRef]
13. Babiak, P.; Hare, R.D.; McLaren, T. *Snakes in Suits: When Psychopaths Go to Work*; Regan Books: New York, NY, USA, 2006.
14. Baggio, M.C.; Benning, S.D. The influence of psychopathic traits and strategic harshness on point gain and cooperation rate in the Prisoner's Dilemma. *Personal. Individ. Differ.* **2022**, *186*, 111344. [CrossRef]
15. Harris, G.T.; Rice, M.E. Treatment of psychopathy. In *Handbook of Psychopathy*; Patrick, C.J., Ed.; The Guilford Press: New York, NY, USA, 2006; pp. 555–572.
16. Stenason, L.; Vernon, P.A. The Dark Triad, reinforcement sensitivity and substance use. *Pers. Individ. Differ.* **2016**, *94*, 59–63. [CrossRef]
17. Ellis, H. Auto-erotism: A psychological study. *Alienist Neurol. (1880–1920)* **1898**, *19*, 260.
18. Campbell, W.K.; Hoffman, B.J.; Campbell, S.M.; Marchisio, G. Narcissism in organizational contexts. *Hum. Resour. Manag. Rev.* **2011**, *21*, 268–284. [CrossRef]
19. Nenadić, I.; Lorenz, C.; Gaser, C. Narcissistic personality traits and prefrontal brain structure. *Sci. Rep.* **2021**, *11*, 15707. [CrossRef]
20. Malesza, M.; Kalinowski, K. Dark triad and impulsivity—An ecological momentary assessment approach. *Curr. Psychol.* **2021**, *40*, 3682–3690. [CrossRef]
21. Maneiro, L.; López-Romero, L.; Gómez-Fraguela, J.A.; Cutrín, O.; Romero, E. Pursuing the dark triad. *J. Individ. Differ.* **2018**, *40*. [CrossRef]
22. Malesza, M. The effects of the Dark Triad traits in prisoner's dilemma game. *Curr. Psychol.* **2020**, *39*, 1055–1062. [CrossRef]
23. Whiteside, S.P.; Lynam, D.R. The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. *Pers. Individ. Differ.* **2001**, *30*, 669–689. [CrossRef]
24. Barratt, E.S. Anxiety and impulsiveness related to psychomotor efficiency. *Percept. Mot. Ski.* **1959**, *9*, 191–198. [CrossRef]
25. Lee, K.; Ashton, M.C. Psychometric Properties of the HEXACO Personality Inventory. *Multivar. Behav. Res.* **2004**, *39*, 329–358. [CrossRef]
26. Ashton, M.C.; Lee, K.; Goldberg, L.R. The IPIP–HEXACO scales: An alternative, public-domain measure of the personality constructs in the HEXACO model. *Pers. Individ. Differ.* **2007**, *42*, 1515–1526. [CrossRef]
27. Ashton, M.C.; Lee, K. On the relations between HEXACO agreeableness (versus anger) and honesty-humility. *Scand. J. Psychol.* **2021**, *62*, 887–894. [CrossRef]
28. De Vries, R.E.; van Kampen, D. The HEXACO and 5DPT models of personality: A comparison and their relationships with psychopathy, egoism, pretentiousness, immorality, and Machiavellianism. *J. Personal. Disord.* **2010**, *24*, 244–257. [CrossRef] [PubMed]
29. Columbus, S. Honesty-Humility, Beliefs, and Prosocial Behaviour: A Test on Stockpiling during the COVID-19 Pandemic. *Collabra Psychol.* **2021**, *7*, 19028. [CrossRef]
30. Deutchman, P.; Sullivan, J. The Dark Triad and framing effects predict selfish behavior in a one-shot Prisoner's Dilemma. *PLoS ONE* **2018**, *13*, e0203891. [CrossRef]
31. Malesza, M. Grandiose narcissism and vulnerable narcissism in prisoner's dilemma game. *Personal. Individ. Differ.* **2020**, *158*, 109841. [CrossRef]
32. Curry, O.; Chesters, M.J.; Viding, E. The psychopath's dilemma: The effects of psychopathic personality traits in one-shot games. *Pers. Individ. Differ.* **2011**, *50*, 804–809. [CrossRef]
33. Hilbig, B.E.; Kieslich, P.J.; Henninger, F.; Thielmann, I.; Zettler, I. Lead Us (Not) into Temptation: Testing the Motivational Mechanisms Linking Honesty–Humility to Cooperation. *Eur. J. Pers.* **2018**, *32*, 116–127. [CrossRef]
34. Zettler, I.; Hilbig, B.E.; Heydasch, T. Two sides of one coin: Honesty–Humility and situational factors mutually shape social dilemma decision making. *J. Res. Pers.* **2013**, *47*, 286–295. [CrossRef]
35. Thielmann, I.; Hilbig, B.E. Is it all about the money? A re-analysis of the link between Honesty-Humility and Dictator Game giving. *J. Res. Pers.* **2018**, *76*, 1–5. [CrossRef]
36. Columbus, S.; Thielmann, I.; Balliet, D. Situational Affordances for Prosocial Behaviour: On the Interaction between Honesty–Humility and (Perceived) Interdependence. *Eur. J. Pers.* **2019**, *33*, 655–673. [CrossRef]
37. Feingold, A. Gender differences in personality: A meta-analysis. *Psychol. Bull.* **1994**, *116*, 429–456. [CrossRef]
38. Costa, P.T., Jr.; Terracciano, A.; McCrae, R.R. Gender differences in personality traits across cultures: Robust and surprising findings. *J. Personal. Soc. Psychol.* **2001**, *81*, 322. [CrossRef]

39. Grijalva, E.; Newman, D.A.; Tay, L.; Donnellan, M.B.; Harms, P.D.; Robins, R.W.; Yan, T. Gender differences in narcissism: A meta-analytic review. *Psychol. Bull.* **2015**, *141*, 261–310. [CrossRef]

40. Muris, P.; Merckelbach, H.; Otgaar, H.; Meijer, E. The malevolent side of human nature: A meta-analysis and critical review of the literature on the dark triad (narcissism, Machiavellianism, and psychopathy). *Perspect. Psychol. Sci.* **2017**, *12*, 183–204. [CrossRef]

41. Cross, C.P.; Copping, L.T.; Campbell, A. Sex differences in impulsivity: A meta-analysis. *Psychol. Bull.* **2011**, *137*, 97–130. [CrossRef]

42. Kivikangas, J.M.; Kätsyri, J.; Järvelä, S.; Ravaja, N. Gender differences in emotional responses to cooperative and competitive game play. *PLoS ONE* **2014**, *9*, e100318. [CrossRef]

43. Dorrough, A.R.; Glöckner, A. A cross-national analysis of sex differences in prisoner's dilemma games. *Br. J. Soc. Psychol.* **2019**, *58*, 225–240. [CrossRef]

44. Mengel, F. Risk and Temptation: A Meta-study on Prisoner's Dilemma Games. *Econ. J.* **2018**, *128*, 3182–3209. [CrossRef]

45. Friedl, A.; Pondorfer, A.; Schmidt, U. Gender differences in social risk taking. *J. Econ. Psychol.* **2020**, *77*, 102182. [CrossRef]

46. Sarin, R.; Wieland, A. Risk aversion for decisions under uncertainty: Are there gender differences? *J. Behav. Exp. Econ.* **2016**, *60*, 1–8. [CrossRef]

47. Charness, G.; Gneezy, U. Strong Evidence for Gender Differences in Risk Taking. *J. Econ. Behav. Organ.* **2012**, *83*, 50–58. [CrossRef]

48. Wood, W.; Eagly, A.H. Biosocial construction of sex differences and similarities in behavior. In *Advances in Experimental Social Psychology*; Academic Press: Cambridge, MA, USA, 2012; Volume 46, pp. 55–123.

49. Babalola, S.S.; Nwanzu, C.L. The moderating role of gender in the effect of self-monitoring personality trait on emotional labor strategies. *Cogent Bus. Manag.* **2022**, *9*, 2046679. [CrossRef]

50. Jonason, P.K.; Webster, G.D. The dirty dozen: A concise measure of the dark triad. *Psychol. Assess.* **2010**, *22*, 420–432. [CrossRef]

51. Czarna, A.Z.; Jonason, P.; Edufner, M.; Kossowska, M. The Dirty Dozen Scale: Validation of a Polish Version and Extension of the Nomological Net. *Front. Psychol.* **2016**, *7*, 445. [CrossRef]

52. Field, A.P. *Discovering Statistics: The Reality Enigma*; SAGE: Thousand Oaks, CA, USA, 2016.

53. Hayes, A.F. Mediation, moderation, and conditional process analysis. In *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach Edn*; Guilford Publications: New York, NY, USA, 2013; Volume 1, p. 20.

54. Peng CY, J.; Lee, K.L.; Ingersoll, G.M. An introduction to logistic regression analysis and reporting. *J. Educ. Res.* **2002**, *96*, 3–14. [CrossRef]

55. Malesza, M.; Ostaszewski, P. Dark side of impulsivity—Associations between the Dark Triad, self-report and behavioral measures of impulsivity. *Personal. Individ. Differ.* **2016**, *88*, 197–201. [CrossRef]

56. Szabó, E.; Jones, D.N. Gender differences moderate Machiavellianism and impulsivity: Implications for Dark Triad research. *Pers. Individ. Differ.* **2019**, *141*, 160–165. [CrossRef]

57. Spector, P.E.; Zhou, Z.E. The Moderating Role of Gender in Relationships of Stressors and Personality with Counterproductive Work Behavior. *J. Bus. Psychol.* **2014**, *29*, 669–681. [CrossRef]

58. Bowling, N.A.; Burns, G.N. Sex as a Moderator of the Relationships between Predictor Variables and Counterproductive Work Behavior. *J. Bus. Psychol.* **2015**, *30*, 193–205. [CrossRef]

59. Chen, Y.; Yan, X.; Fan, W.; Gordon, M. The joint moderating role of trust propensity and gender on consumers' online shopping behavior. *Comput. Hum. Behav.* **2015**, *43*, 272–283. [CrossRef]

60. Tarka, P.; Kukar-Kinney, M.; Harnish, R.J. Consumers' personality and compulsive buying behavior: The role of hedonistic shopping experiences and gender in mediating-moderating relationships. *J. Retail. Consum. Serv.* **2022**, *64*, 102802. [CrossRef]

61. Paulhus, D.L.; Vazire, S. The self-report method. *Handb. Res. Methods Personal. Psychol.* **2007**, *1*, 224–239.