Supplemental Information

Associations of Aversive ('Dark') Traits and Affiliative ('Light') Traits with Moral-Dilemma Judgments: A Preregistered Exploratory Analysis Using the CNI Model

Nyx L. Ng[1], Craig S. Neumann[2], Dillon M. Luke[1], Bertram Gawronski[1]

[1] University of Texas at Austin

[2] University of North Texas

Model Fit of Path Analysis

To provide a test of model fit and avoid under-identified models, we excluded several scales that showed little association with the CNI model parameters in the regression analyses in follow-up trimmed path analyses. In the trimmed path analyses with the superordinate SD4 and LTS factors as predictor variables, we excluded the paths between both the superordinate aversive and affiliative traits and the *C* and *I* parameters (Figure S1). In the trimmed path analyses with the subordinate SD4 and LTS factors as predictor variables, we excluded the following paths: (1) the paths between the *C* parameter on sensitivity to consequences and both Machiavellianism and humanism, (2) the path between the *N* parameter and faith in humanity, and (3) the paths between the *I* parameter on general preference for inaction over action and both Machiavellianism and psychopathy (Figure S2). Fit was good for the superordinate model, $\chi^2(4) = 6.34$, $p = .175$, CFI = .95, RMSEA = .03, RMSEA 90% CI [.00, .07], and subordinate model, $\chi^2(5) = 1.42$, $p = .922$, CFI = 1.00, RMSEA = .00, RMSEA 90% CI [.00, .02]. Omitting the identified variables in the trimmed path models also did not alter the substantive pattern of results compared to the models including all variables as predictors.

Robustness Check

To rule out undue effects of potential outliers, we exploratorily re-ran all analyses after outlier exclusion. Toward this end, we first computed the median absolute deviation for each aversive and affiliative trait, after which we applied a moderately conservative criterion in which cases are identified as outliers if they lie beyond 2.5 absolute deviations away from the median (Leys et al., 2013). A total of 121 cases were identified as potential outliers in at least one of the seven SD4 and LTS subscales used in the study, with psychopathy ($n = 33$) and Kantianism ($n = 74$) being the subordinate aversive and affiliative traits with the highest numbers of outliers, respectively. The reduced sample thus included 555 cases. Excluding the potential outliers led to a reduction in internal consistency across the subscales, with the subscales of Machiavellianism (α = .51, mean inter-item correlation = .13), psychopathy (α = .57, mean inter-item correlation = .18), and particularly Kantianism (α = .21, mean inter-item correlation = .07) showing poor reliability after outlier exclusion (Table S1).

After outlier-exclusion, the following differences in associations were found in the correlational analyses (Table S2): (1) the traditional score was no longer significantly associated with the superordinate affiliative trait, (2) narcissism is now positively associated with superordinate affiliative trait, and (3) the negative association between psychopathy and sensitivity to moral norms is no longer statistically significant. Consistent with the original multiple-regression analyses, sadism emerged as the only significant subordinate predictor of sensitivity to consequences, and none of the traits, both in the superordinate and subordinate factor models, predicted general preference for inaction over action. Likewise, both the superordinate aversive and affiliative factors significantly predicted sensitivity to moral norms in opposing ways, and the negative association between sadism and sensitivity to moral norms also remained statistically significant (Table S3).

Inconsistent with the multiple-regression and path analyses (Figures S3 and S4) using the full sample, the reanalyses yielded the following results: (1) the previously marginal negative association between Machiavellianism and sensitivity to moral norms is now statistically significant; (2) the previously marginal association between Kantianism and sensitivity to moral norms is now statistically significant; and (3) the association between humanism and sensitivity to moral norms is now marginal. The former two findings make intuitive sense given that Machiavellianism is associated with tendencies to disregard conventional morality (Muris et al.,

2017) and Kantianism is directly related to tendencies to endorse deontological principles (Kaufman et al., 2019). However, caution should be taken when interpreting these associations because both the Kantianism and Machiavellianism subscales had poor internal consistency after outlier exclusion.

CAN Algorithm

In the CNI model, the *I* parameter on general preference for inaction over action is positioned at the bottom of the processing tree for methodological reasons (see Gawronski et al., 2017, 2020), which results in estimations of the *I* parameter having greater proportions of measurement error than the *C* and *N* parameters on sensitivity to consequences and moral norms. The null associations found between all traits and the CNI model's *I* parameter could have thus been an artifact of the CNI model's hierarchical structure.

To address this possibility, we re-analyzed the relations using the CAN algorithm, which algebraically calculates the three model parameters concurrently rather than hierarchically (see Liu & Liao, 2021). Unlike the CNI model, the CAN algorithm estimates an index (*A*) that reflects the probability of action (versus inaction) responses to all four dilemma variants and is thus directionally opposite to the CNI models' *I* parameter. Also, unlike the CNI model which produces parameter estimates with low intercorrelations, the CAN algorithm's concurrent calculation of the three factors often leads to high correlations between the three parameters. High correlations between the three parameters can lead to false-positives or false-negatives when their shared variance is not accounted for in statistical models. Thus, in the multiple regression analyses predicting each of the three CAN algorithm parameters, we controlled for the effects of the other two parameters. Internal consistency estimates of the CAN algorithm parameters are as follows: the *N* parameter on sensitivity to moral norms had the highest consistency (Cronbach's α = .71, mean inter-item r = .55), followed by the *C* parameter on consequences (Cronbach's α = .57, mean inter-item r = .43), and the *A* parameter on action preferences (Cronbach's α = .33, mean inter-item r = .20). Results of the zero-order correlational analyses are presented in Table S4, and the multiple regression analyses are presented in Table S5. Results of the path analyses are presented in Figures S5 and S6.

***C* parameter**

Consistent with the CNI model analyses, both the correlational and multiple regression analyses suggest no associations between the superordinate affiliative and aversive traits and the *C* parameter on sensitivity to consequences. The significant positive association between the CAN algorithm's *C* parameter emerged in the zero-order correlational analyses and path analyses controlling for covariances between predictor and criterion variables. However, the relationship between sadism and sensitivity to consequences was not statistically significant in the regression model regressing the *C* parameter on the seven traits, *N* parameter, and *A* parameter. Note that sadism significantly predicts sensitivity to consequences when the *N* parameter was not controlled for in the regression model. The relationship between the *C* parameter and sadism thus does not appear to be robust across all analytic approaches.

***N* parameter**

Consistent with the CNI model analyses, the superordinate affiliative trait is positively associated with the CAN algorithm's *N* parameter on sensitivity to moral norms whereas the superordinate aversive trait is negatively associated with it in the correlational analyses. This finding was robust across analytical approaches. The correlations between the CAN algorithm's sensitivity to moral norms and the seven subordinate traits are also consistent with the correlational analyses conducted with the CNI model parameters. Also consistent with the CNI model analyses, the subordinate trait multiple regression analyses found that sadism negatively predicted sensitivity to moral norms. However, inconsistent with the CNI model analyses, Machiavellianism and Kantianism predicted sensitivity to moral norms negatively and positively, respectively. Yet, in the path analyses, only Machiavellianism and sadism emerged as significant predictors of sensitivity to moral norms.

A parameter

Contrary to the findings of the CNI model analyses, correlating the CAN algorithm's *A* parameter on preference for action over inaction with the traits suggests that preference for action is positively associated with the affiliative traits, both on a superordinate and subordinate level. The multiple regression analyses also yielded a positive association between the *A* parameter and superordinate affiliative trait but did not reveal any significant associations between the *A* parameter and the subordinate aversive and affiliative traits. Controlling for the covariances between the three CAN algorithm parameters, the path analyses suggest that the superordinate affiliative trait is positively associated with action preferences such that individuals with elevated affiliative traits tend to respond with action in moral dilemmas. The individual subordinate affiliative traits, however, were not significantly related to the CAN algorithm's *A* parameter in the path analysis with the subordinate traits as predictor variables.

Latent Profile Analysis

In addition to the variable-centered analyses, we conducted an exploratory person-centered latent profile analysis (LPA) of the LTS and SD4 scales with the goal of replicating the subtypes reported by Neumann et al. (2020). We anticipated a 3-class solution with (1) a prosocial subtype profile displaying elevated LTS traits and relatively lower SD4 traits, (2) an antisocial subtype profile displaying the opposite profile, and (3) a middle subtype that would thread between the two previous subtypes with moderate levels of LTS and SD4 traits. M*plus* was used to conduct the LPA with maximum likelihood estimation.

LPA is a variant of finite-mixture modeling that identifies nominal variables underlying continuous data and classifies individuals who are similar on the indicators into latent classes (Hallquist & Wright, 2014; Vermunt & Magidson, 2002). The Bayesian Information Criterion (BIC) and sample-size adjusted BIC are considered reliable indices for selecting the optimal model (Nylund et al., 2007). Models with lower BIC values are preferred. Theoretical considerations and classification accuracy are also useful for selecting optimal models (Neumann et al., 2020). Viable LPA solutions are obtained when the average latent class probabilities (accuracy) for the most likely class membership are greater than .80 (Mokros et al., 2015). Monte Carlo simulations indicate larger samples (>250), more (versus fewer) indicators, and greater degree of class separation (large effect size) influence the likelihood of uncovering true latent class solutions (Tein et al., 2013). The current study has a sample size of 676 and seven indicators, and class separation was expected to be large (partial $\eta^2 > .20$). The LPA was conducted with this sample aggregating across gender. Class separation effect size was assessed via MANOVA with latent class as the independent variable and LTS/SD4 scales as dependent variables. Next, we conducted chi-square analyses on gender, ethnicity, and age to assess demographics associated with the subtype profiles. Demographic variables that have significantly different proportions across the subtype profiles were included as covariates in the one-way MANOVA assessing how the subtypes differed on the CNI model parameters.

The LPA results indicated that multiple latent classes could be extracted from the sample using the LTS and SD4 scales. The adjusted BIC values dropped greatly from the 1-class ($BIC_{adj.}$ = 8136.52 and 2-class solutions ($BIC_{adj.}$ = 7681.79) to the 3-class solution ($BIC_{adj.}$ = 7541.55).

The 3-class solution also resulted in high classification accuracy (91.4%) with latent classes consistent with previous research (Neumann et al., 2020). The adjusted BIC value of the 4-class solution was smaller than the 3-class solution ($BIC_{adj.}$ = 7379.72), but one subtype represented less than 5% of the total sample and there was a drop in classification accuracy (86.1%) as compared to the 3-class solution. The 4-class solution was thus not viable. Entering the LPA subtypes as a predictor of the superordinate LTS and SD4 scales indicated that class separation for the 3-class solution on the LTS and SD4 scales was strong ($\eta_p^2 > .26$).

The pattern of LTS and SD4 profiles for the three subtypes is displayed in Figure S7. Consistent with prior research (Neumann et al., 2020), the Prosocial subtype contained the highest proportion of cases (66.4%) whereas the Antisocial subtype had the lowest proportion of cases (5.5%). There were proportionally more women (70.6%) than men in the Prosocial subtype, more men (67.6%) than women in the Antisocial subtype, and more women than men in the Middle subtype (67.2%). The differences in the proportions of men versus women[1] were significant across subtypes, $\chi^2(2) = 22.74$, $p < .001$, but the Prosocial and Middle subtypes did not significantly differ in their proportions of men and women, $\chi^2(2) = 3.09$, $p = .213$. For ethnicity categories with adequate numbers of representation, there were no significant differences in ethnicity by subtype, *p*s > .05. Finally, consistent with previous research (Neumann et al., 2020), there was a trend for differences in age across subtypes, with the Prosocial subtype the oldest ($M = 43.90$, $SD = 12.42$), followed by the Middle subtype ($M = 42.39$, $SD = 11.66$), and then the Antisocial subtype ($M = 39.92$, $SD = 12.30$), but these age differences were not statistically significant, $F(2, 673) = 2.49$, $p = .084$, $\eta_p^2 = .01$.

The results of a one-way MANOVA with the CNI model parameters as dependent variables and gender as a covariate indicated that the main effect for Subtype was significant,

$F(6, 1338) = 4.53$, $p < .001$, $\eta_p^2 = .02$, Wilk's $\lambda = .96$, with only mean scores on the *N* parameter on sensitivity to moral norms being significantly different across Subtypes, $F(2, 671) = 11.61$, $p < .001$, $\eta_p^2 = .03$. Planned comparisons of the Subtype main effect with Bonferroni correction suggest a significant difference in norm sensitivity between the Antisocial subtype (*p*s < .001) and both Middle and Prosocial subtypes, the latter of which also significantly differed ($p = .049$). Figure S8 shows the pattern of differences on the CNI model parameters for the three subtypes. In sum, these findings suggest that the three subtypes did not differ in terms of their sensitivity to consequences or general action tendencies but differed in their sensitivity to moral norms of harm.

References

Gawronski, B., Armstrong, J., Conway, P., Friesdorf, R., & Hütter, M. (2017). Consequences, norms, and generalized inaction in moral dilemmas: The CNI model of moral decision-making. *Journal of Personality and Social Psychology*, *113*(3), 343–376. https://doi.org/10.1037/pspa0000086

Gawronski, B., Conway, P., Hütter, M., Luke, D. M., Armstrong, J., & Friesdorf, R. (2020). On the validity of the CNI model of moral decision-making: Reply to Baron and Goodwin (2020). *Judgment and Decision Making*, *15*(6), 1054–1072.

Hallquist, M. N., & Wright, A. G. C. (2014). Mixture modeling methods for the assessment of normal and abnormal personality, part I: Cross-sectional models. *Journal of Personality Assessment*, *96*(3), 256–268. https://doi.org/10.1080/00223891.2013.845201

Horsten, L. K., Moshagen, M., & Hilbig, B. E. (2023). On the (number of) aversive traits it takes to approximate the common core of aversive personality. *Scientific Reports*, *13*(1), 15021. https://doi.org/10.1038/s41598-023-42115-z

Kaufman, S. B., Yaden, D. B., Hyde, E., & Tsukayama, E. (2019). The Light vs. Dark Triad of personality: Contrasting two very different profiles of human nature. *Frontiers in Psychology*, *10*, 467. https://doi.org/10.3389/fpsyg.2019.00467

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Liu, C., & Liao, J. (2021). CAN algorithm: An individual level approach to identify consequence and norm sensitivities and overall action/inaction preferences in moral decision-making. *Frontiers in Psychology*, *11*, 547916. https://doi.org/10.3389/fpsyg.2020.547916

Mokros, A., Hare, R. D., Neumann, C. S., Santtila, P., Habermeyer, E., & Nitschke, J. (2015). Variants of psychopathy in adult male offenders: A latent profile analysis. *Journal of Abnormal Psychology*, *124*(2), 372–386. https://doi.org/10.1037/abn0000042

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*(2), 183–204. https://doi.org/10.1177/1745691616666070

Neumann, C. S., Kaufman, S. B., ten Brinke, L., Yaden, D. B., Hyde, E., & Tsykayama, E. (2020). Light and dark trait subtypes of human personality – A multi-study person-centered approach. *Personality and Individual Differences*, *164*, 110121. https://doi.org/10.1016/j.paid.2020.110121

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(4), 535–569. https://doi.org/10.1080/10705510701575396

Tein, J.-Y., Coxe, S., & Cham, H. (2013). Statistical power to detect the correct number of classes in latent profile analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(4), 640–657. https://doi.org/10.1080/10705511.2013.824781

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. In *Applied Latent Class Analysis* (pp. 89–106). Cambridge University Press.

Footnotes

[1]We only included gender identifications as *male* and *female* but not *other* due to the low number of observations in the latter.

Table S1. *Internal consistency estimates of the SD4 measure of aversive traits, LTS measure of affiliative traits, traditional dilemma variant, and CNI model parameters after outlier exclusion*

	After outlier exclusion	
Variables	α	Mean inter-item correlation
Aversive traits	.76	.10
Machiavellianism	.51	.13
Narcissism	.73	.28
Psychopathy	.57	.18
Sadism	.62	.19
Affiliative traits	.70	.16
Faith in humanity	.71	.39
Humanism	.64	.31
Kantianism	.21	.07
Traditional score	.62	.14
CNI model parameters		
C parameter	.56	.42
N parameter	.54	.38
I parameter	.25	.15

Note. *C* parameter = sensitivity to consequences; *N* parameter = sensitivity to moral norms; *I* parameter = general preference for inaction over action.

Table S2. *Zero-order correlations between study variables after outlier exclusion*

	1	2	3	4	5	6	7	8	9	10	11	12
1. Traditional score	1											
2. C parameter	.71***	1										
3. N parameter	-.62***	-.06	1									
4. I parameter	-.30***	.08	.28***	1								
5. Aversive trait	.13**	.01	-.15***	-.02	1							
6. Affiliative trait	-.07	-.01	.13**	-.03	-.14***	1						
7. Machiavellianism	.08*	-.02	-.12**	-.03	.54***	-.03	1					
8. Narcissism	.00	-.05	-.03	-.01	.70***	.10*	.17***	1				
9. Psychopathy	.07	-.02	-.07	-.01	.70***	-.19***	.14**	.39***	1			
10. Sadism	.19***	.12**	-.18***	-.00	.69***	-.28***	.26***	.17***	.35***	1		
11. Faith in humanity	-.04	-.01	.05	-.06	-.08	.82***	-.06	.15***	-.11**	-.21***	1	
12. Humanism	-.09*	-.04	.11**	-.04	-.04	.87***	.04	.20***	-.13**	-.23***	.67***	1
13. Kantianism	-.03	.06	.15***	.05	-.25***	.55***	-.07	-.21***	-.19***	-.17***	.15***	.26***

Note. *$p < .05$, **$p < .01$, ***$p < .001$. *C* parameter = sensitivity to consequences; *N* parameter = sensitivity to moral norms; *I* parameter = general preference for inaction over action.

Table S3. *Results of multiple-regression analyses regressing CNI model parameters on aversive and affiliative traits after outlier exclusion*

	C parameter			*N* parameter			*I* parameter		
Variable	B	95% CI	Adj. R^2	B	95% CI	Adj. R^2	B	95% CI	Adj. R^2
Superordinate factors			-.00			.03			-.00
Aversive trait	0.01	[-0.04, 0.05]		-0.12**	[-0.19, -0.05]		-0.02	[-0.09, 0.05]	
Affiliative trait	-0.00	[-0.04, 0.03]		0.08**	[0.02, 0.14]		-0.03	[-0.09, 0.04]	
Subordinate factors			.01			.05			-.00
Machiavellianism	-0.02	[-0.05, 0.02]		-0.06*	[-0.12, -0.00]		-0.02	[-0.08, 0.04]	
Narcissism	-0.01	[-0.04, 0.02]		0.00	[-0.04, 0.05]		0.01	[-0.04, 0.06]	
Psychopathy	-0.02	[-0.05, 0.02]		0.01	[-0.05, 0.07]		-0.00	[-0.07, 0.06]	
Sadism	0.05***	[0.02, 0.08]		-0.07**	[-0.12, -0.02]		-0.00	[-0.05, 0.05]	
Faith in humanity	0.01	[-0.02, 0.04]		-0.03	[-0.08, 0.01]		-0.03	[-0.08, 0.02]	
Humanism	-0.02	[-0.05, 0.02]		0.06†	[-0.00, 0.13]		-0.01	[-0.08, 0.06]	
Kantianism	0.03	[-0.01, 0.06]		0.07*	[0.01, 0.13]		0.04	[-0.02, 0.11]	

Note. †$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$. *C* parameter = sensitivity to consequences; *N* parameter = sensitivity to moral norms; *I* parameter = general preference for inaction over action.

Table S4. *Zero-order correlations between study variables and CAN algorithm parameters*

	1	2	3	4	5	6	7	8	9	10	11
1. *C* parameter	1										
2. *N* parameter	-.47***	1									
3. *A* parameter	.02	.09*	1								
4. Aversive trait	.05	-.23***	.02	1							
5. Affiliative trait	-.00	.16***	.11**	-.24***	1						
6. Machiavellianism	.04	-.16***	.01	.59***	-.18***	1					
7. Narcissism	-.04	-.04	.02	.67***	.05	.22***	1				
8. Psychopathy	.01	-.17***	.02	.75***	-.24***	.21***	.38***	1			
9. Sadism	.13***	-.28***	.01	.73***	-.31***	.30***	.19***	.47***	1		
10. Faith in humanity	.01	.09*	.08*	-.16***	.82***	-.14***	.12**	-.18***	-.24***	1	
11. Humanism	-.03	.15***	.11**	-.15***	.86***	-.09*	.15***	-.18***	-.29***	.69***	1
12. Kantianism	.02	.14***	.08*	-.28***	.67***	-.23***	-.18***	-.18***	-.18***	.26***	.35***

Note. **p* < .05, ***p* < .01, ****p* < .001. *C* parameter = sensitivity to consequences; *N* parameter = sensitivity to moral norms; *A* parameter = probability of action (versus inaction) responses.

Table S5. *Results of multiple-regression analyses regressing CAN algorithm parameters on aversive and affiliative traits*

	C parameter			*N* parameter			*A* parameter		
Variable	B	95% CI	Adj. R^2	B	95% CI	Adj. R^2	B	95% CI	Adj. R^2
Superordinate model			.23			.28			.02
Aversive trait	-.02	[-0.05, 0.01]		-0.12***	[-0.16, -0.08]		0.01†	[-0.00, 0.02]	
Affiliative trait	.02†	[-0.00, 0.05]		0.06**	[0.02, 0.09]		0.01**	[0.00, 0.02]	
CAN parameters									
C parameter				-0.70***	[-0.80, -0.60]		0.02	[-0.01, 0.06]	
N parameter	-0.33***	[-.37, -.28]					0.03**	[0.01, 0.05]	
A parameter	0.15	[-.04, .34]		0.37**	[0.10, 0.64]				
Subordinate model			.23			.29			.01
Machiavellianism	-0.01	[-0.03, 0.04]		-0.04*	[-0.07, 0.00]		0.00	[-0.01, 0.01]	
Narcissism	-0.01	[-0.03, 0.01]		0.01	[-0.02, 0.04]		-0.00	[-0.01, 0.01]	
Psychopathy	-0.02†	[-0.05, 0.01]		-0.03†	[-0.07, 0.00]		0.01	[-0.00, 0.02]	
Sadism	0.02	[-0.01, 0.07]		-0.06***	[-0.09, -0.03]		0.00	[-0.00, 0.01]	
Faith in humanity	0.01	[-0.01, 0.04]		-0.01	[-0.05, 0.02]		0.00	[-0.01, 0.01]	
Humanism	-0.00	[-0.03, 0.02]		0.03	[-0.01, 0.07]		0.01	[-0.00, 0.02]	
Kantianism	0.02†	[-0.00, 0.04]		0.03*	[0.00, 0.06]		0.00	[-0.00, 0.01]	
CAN parameters									
C parameter				-0.68***	[-0.78, -0.58]		0.02	[-0.01, 0.06]	
N parameter	-0.32***	[-0.36, -0.27]					0.03**	[0.01, 0.05]	
A parameter	0.15	[-0.04, 0.33]		0.38**	[0.11, 0.65]				

Note. †$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$. *C* parameter = sensitivity to consequences; *N* parameter = sensitivity to moral norms; *A* parameter = probability of action (versus inaction) responses.

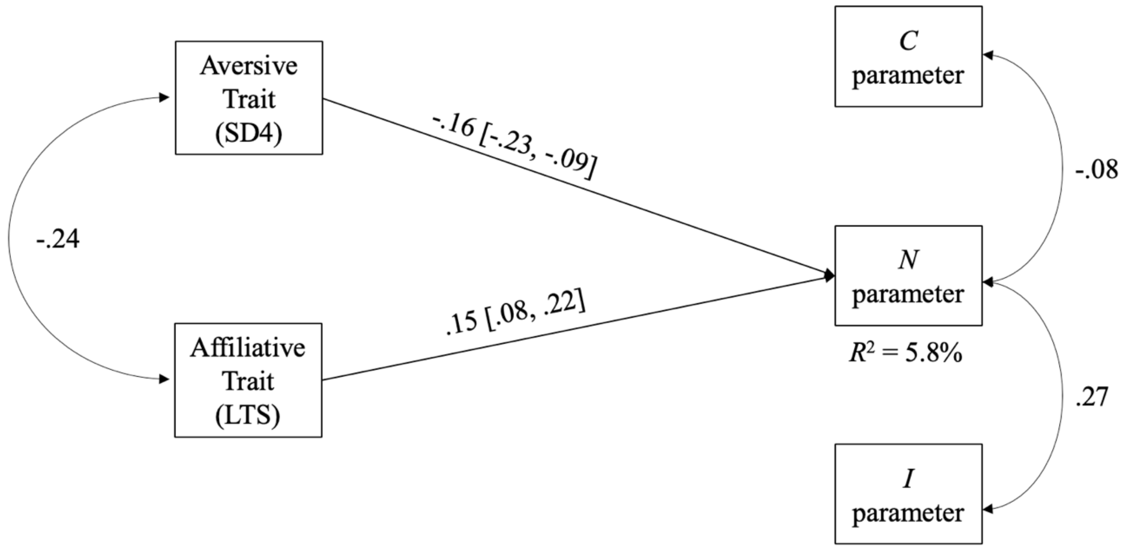

Figure S1. Results of the trimmed path analysis with the superordinate SD4 and LTS traits as predictor variables and CNI model parameters as criterion variables (standardized parameters) with 95% confidence intervals in parentheses. Non-significant paths where $p \geq .05$ are omitted. *C* parameter = sensitivity to consequences; *N* parameter = sensitivity to moral norms; *I* parameter = general preference for inaction over action.

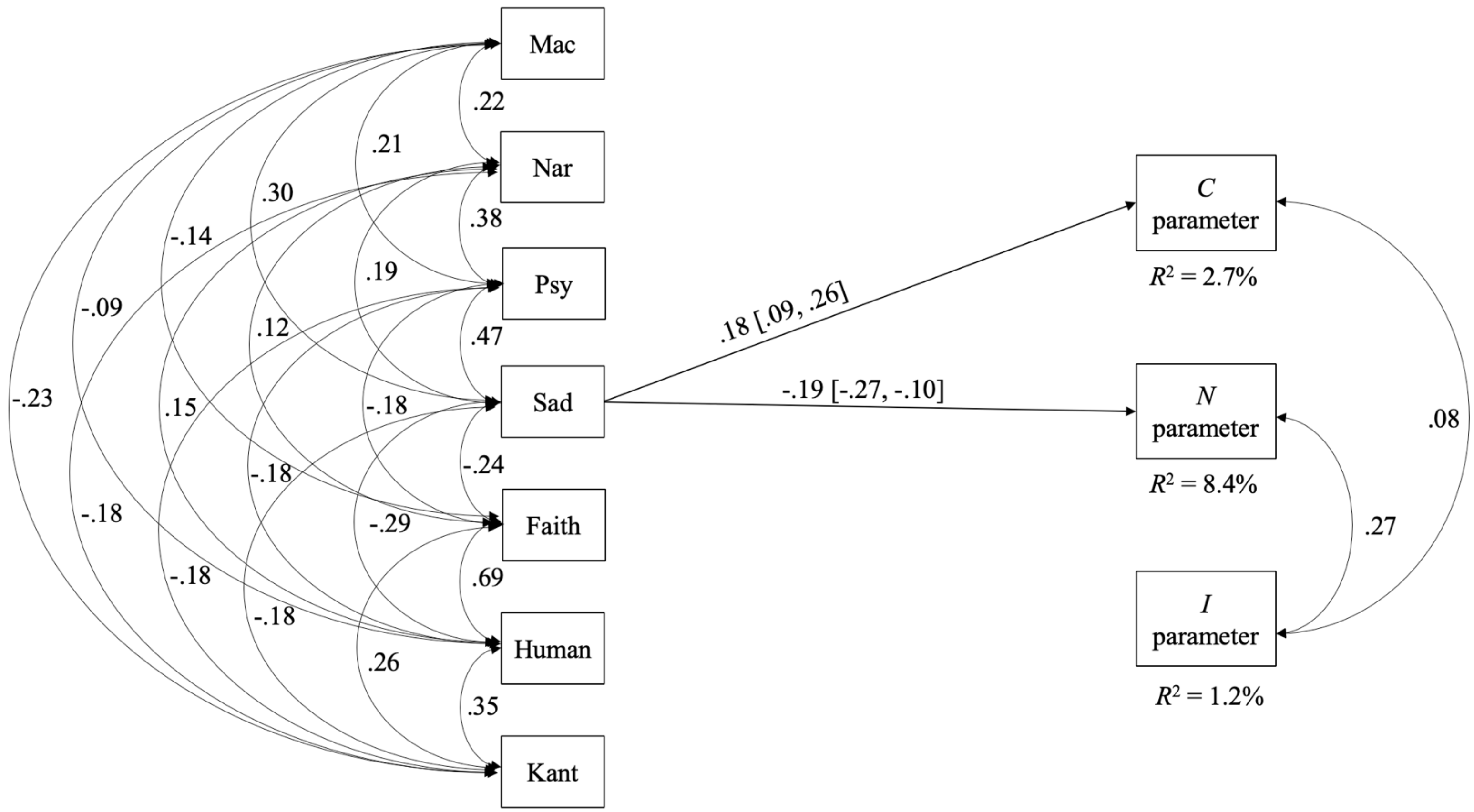

Figure S2. Results of the trimmed path analysis with the subordinate SD4 and LTS traits as predictor variables and CNI model parameters as criterion variables (standardized parameters) with 95% confidence intervals in parentheses. Non-significant paths where $p \geq .05$ are omitted. Mac = Machiavellianism; Nar = Narcissism; Psy = Psychopathy; Sad = Sadism; Faith = Faith in humanity; Human = Humanism; Kant = Kantianism. *C* parameter = sensitivity to consequences; *N* parameter = sensitivity to moral norms; *I* parameter = general preference for inaction over action.

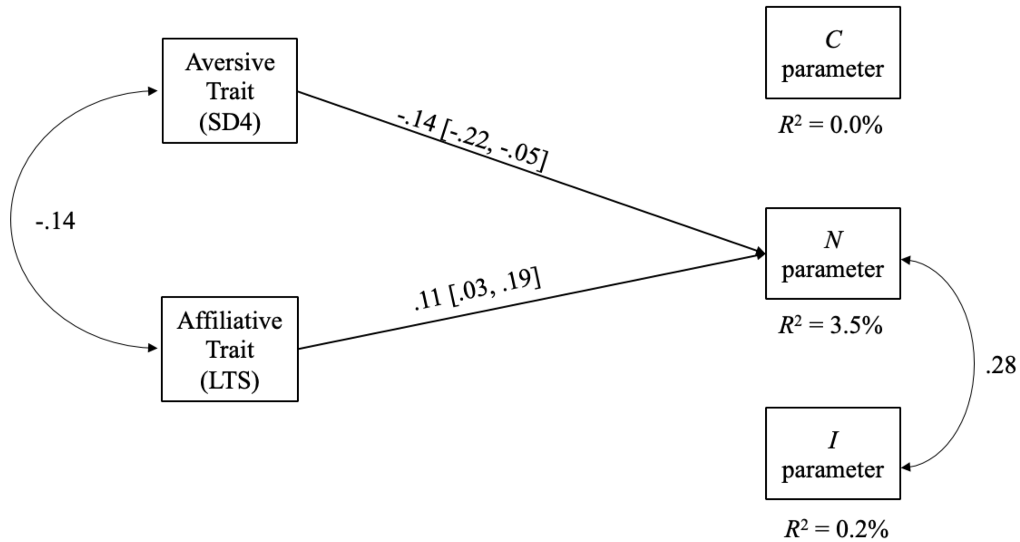

Figure S3. Results of the path analysis after outlier exclusion, with the superordinate aversive (SD4) and affiliative (LTS) traits as predictor variables and CNI model parameters as criterion variables (standardized parameters) with 95% confidence intervals in parentheses. Non-significant paths where $p \geq .05$ are omitted. *C* parameter = sensitivity to consequences; *N* parameter = sensitivity to moral norms; *I* parameter = general preference for inaction over action.

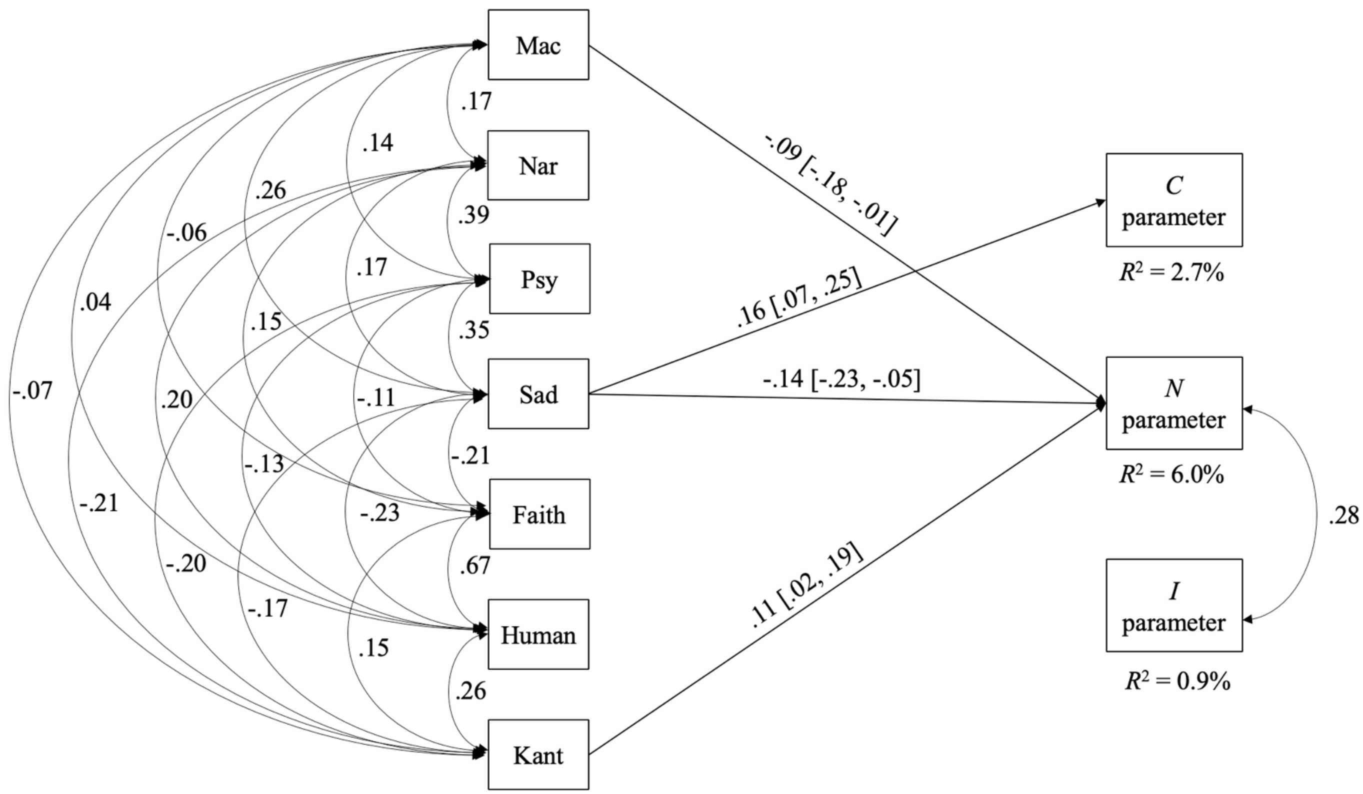

Figure S4. Results of the path analysis after outlier exclusion, with the subordinate SD4 and LTS traits as predictor variables and CNI model parameters as criterion variables (standardized parameters) with 95% confidence intervals in parentheses. Non-significant paths where $p \geq .05$ are omitted. Mac = Machiavellianism; Nar = Narcissism; Psy = Psychopathy; Sad = Sadism; Faith = Faith in humanity; Human = Humanism; Kant = Kantianism. *C* parameter = sensitivity to consequences; *N* parameter = sensitivity to moral norms; *I* parameter = general preference for inaction over action.

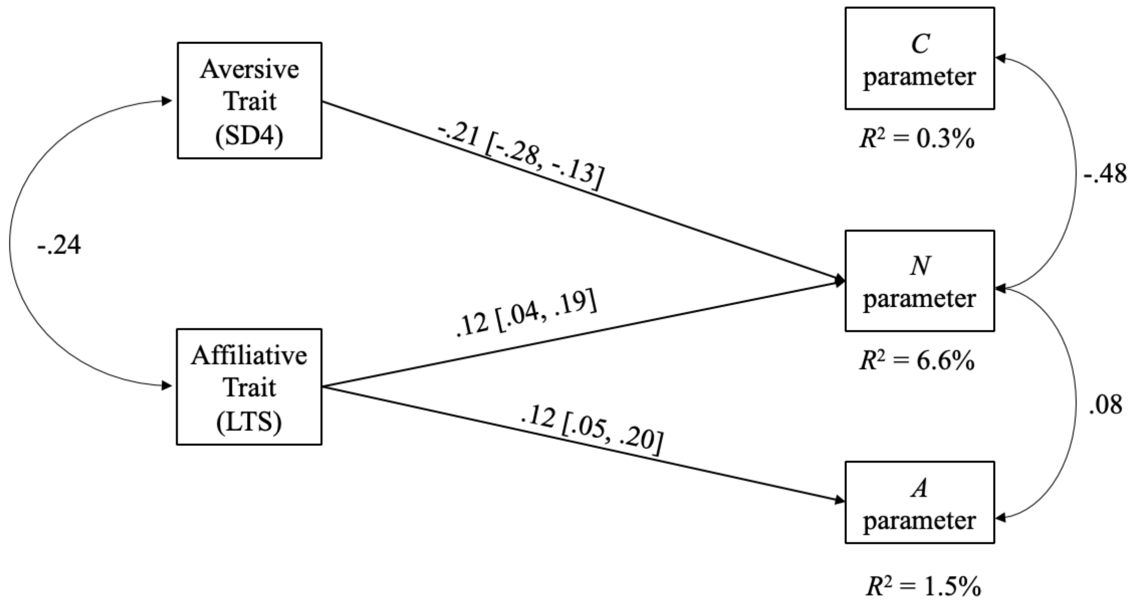

Figure S5. Results of the path analysis with the superordinate SD4 and LTS scales as predictor variables and CAN algorithm parameters as criterion variables (standardized parameters) with 95% confidence intervals in parentheses. Non-significant paths where $p \geq .05$ are omitted. *C* parameter = sensitivity to consequences; *N* parameter = sensitivity to moral norms; *A* parameter = probability of action (versus inaction) responses.

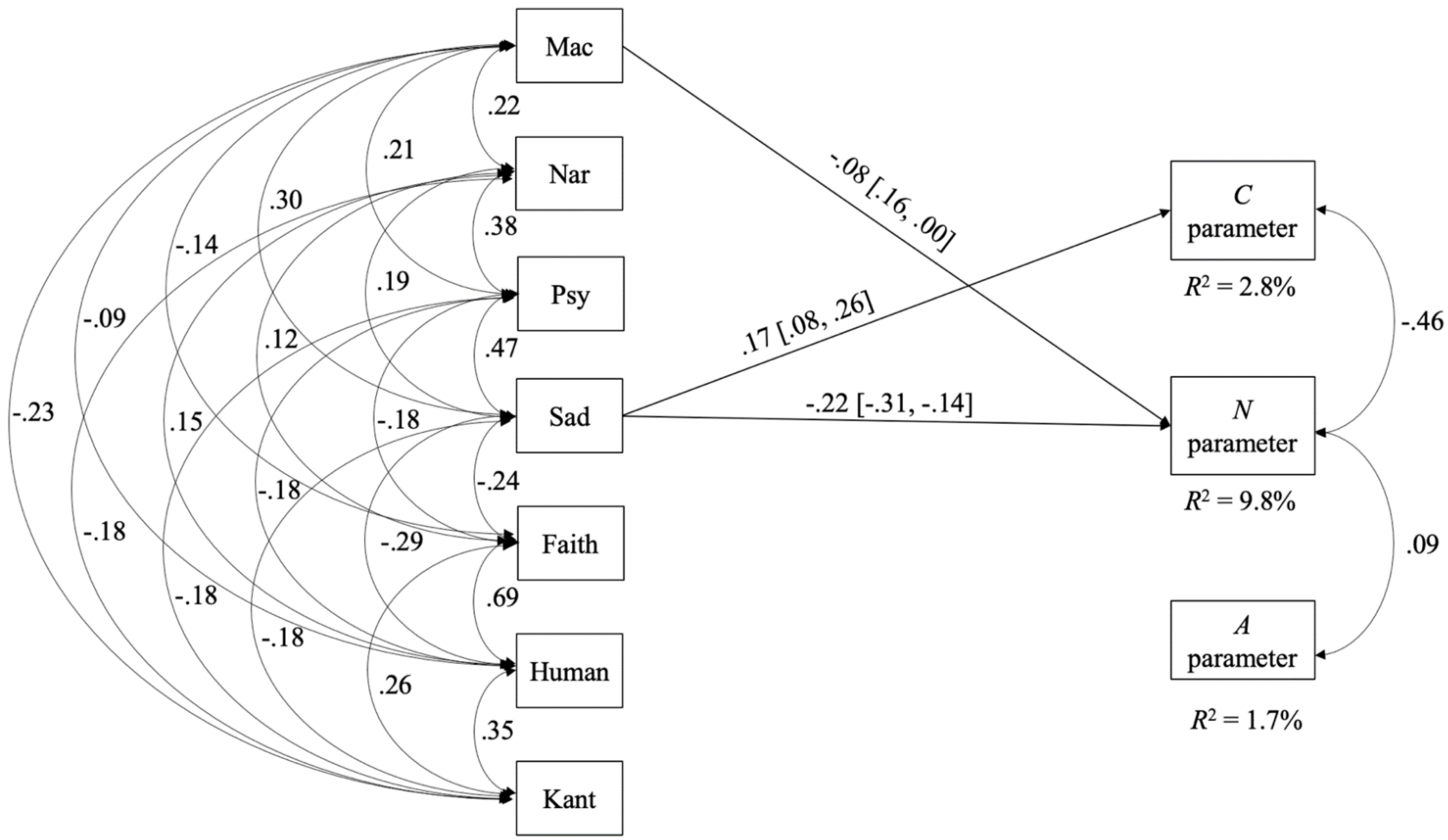

Figure S6. Results of the path analysis with the subordinate SD4 and LTS scales as predictor variables and CAN algorithm parameters as criterion variables (standardized parameters) with 95% confidence intervals in parentheses. Non-significant paths where $p \geq .05$ are omitted. Mac = Machiavellianism; Nar = Narcissism; Psy = Psychopathy; Sad = Sadism; Faith = Faith in humanity; Human = Humanism; Kant = Kantianism. *C* parameter = sensitivity to consequences; *N* parameter = sensitivity to moral norms; *A* parameter = probability of action (versus inaction) responses.

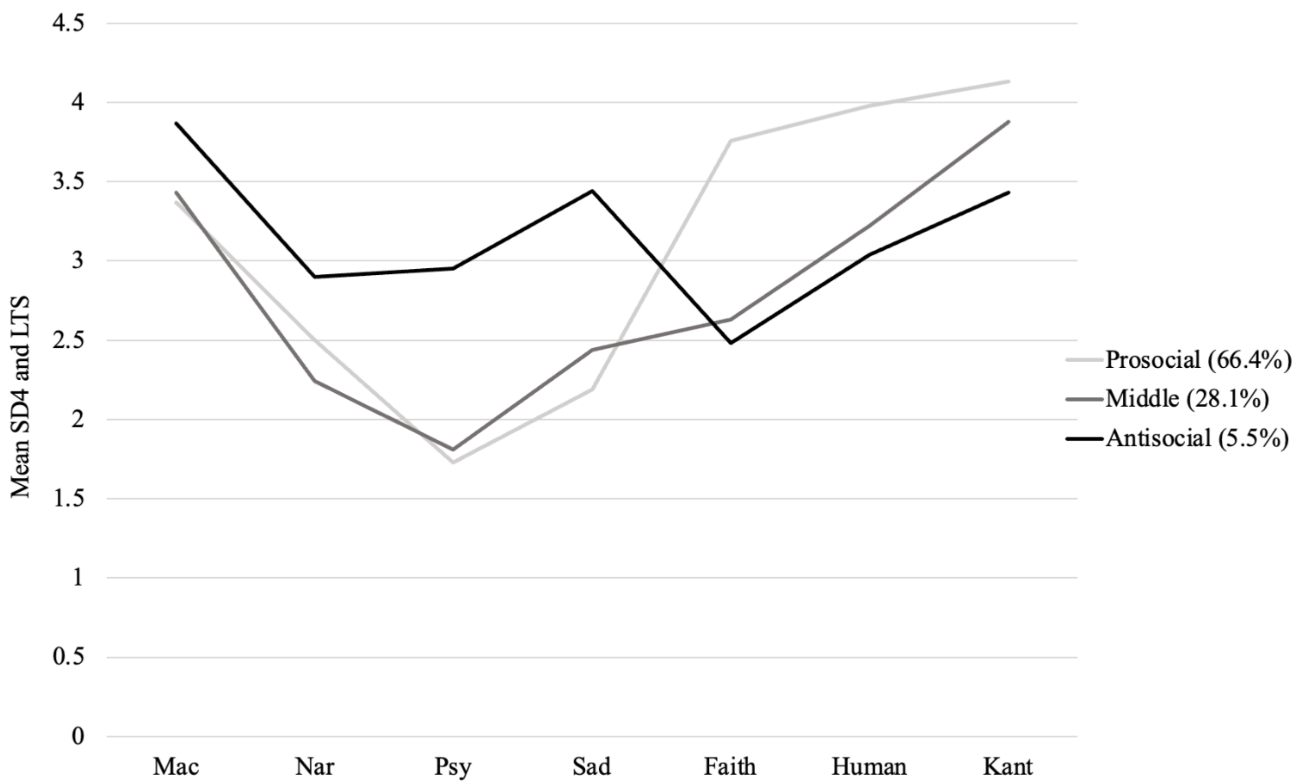

Figure S7. Person-centered latent profile analysis revealed a Prosocial, Middle, and Antisocial subtype. Mac = Machiavellianism; Nar = Narcissism; Psy = Psychopathy; Sad = Sadism; Faith = Faith in Humanity; Human = Humanism; Kant = Kantianism.

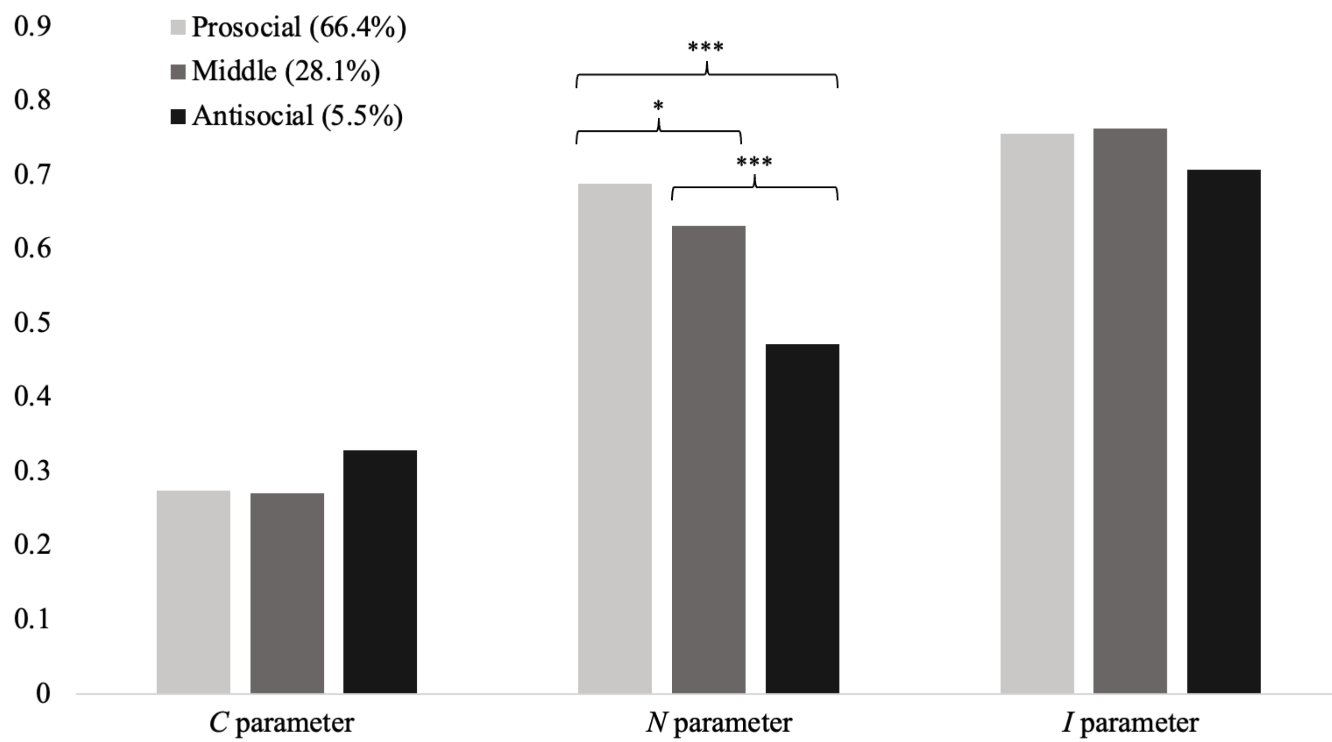

Figure S8. Sensitivity to moral norms discriminate the three subtypes.