Appalachian
STATE UNIVERSITY

Department of Economics Working Paper

Number 23-04 | June 2023

Dark versus Light Personality Types and Moral Choice

David L. Dickinson
Appalachian State University

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-2148
Fax: (828) 262-6105
www.business.appstate.edu/economics

Dark versus light personality types and moral choice

David L Dickinson

Abstract

Dark personality traits have been linked to behaviors commonly understood as unethical, such as fraud, bribe-taking, and marital infidelity. Presumably, more "light" personality traits may be associated with lesser tendencies to be unethical, but many individuals also possess both light and dark trait characteristics. This paper reports results from a preregistered study of over 2400 participants who completed validated short-form personality instruments to assess dark and light personality trait measures—the dark tetrad and a light "triad" of 3 personality dimensions were measured. Furthermore, participants completed 3 tasks of interest that contribute to an understanding or one's ethics: a task assessing prosociality, a task that presents a monetary temptation to be dishonest, and a hypothetical moral dilemma task. The results overall support the hypotheses that dark personality traits predict lower levels of prosociality, higher likelihood of dishonesty, and an increased willingness to make immoral choices overall. Potential mechanisms and implications are examined.

Key Words: Ethics, dark personality, moral choice, experiments.

Acknowledgements: The author is grateful to Appalachian State University and the Walker College of Business for funding the data collection for this project.

1. INTRO

A cluster of dark personality traits named the "dark tetrad" (psychopathy, narcissism, Machiavellianism, sadism) have attracted the interest of researchers interested in understanding personality and decision making. More recently, these have been put in direct contrast with a novel set of "light" personality traits. It is noteworthy that many individuals display dimensions of both dark and light personality traits, such that one's net "lightness" may hold predictive validity regarding decision making. Dark personality traits have been linked behaviors commonly understood as unethical, such as fraud, bribe-taking, and marital infidelity. Presumably, light personality traits would be negatively associated with unethical behaviors, but the balance of these traits in one's personality may be key towards a better understanding of ethical decision tendencies.

This paper reports results from a preregistered study of over 2400 participants who completed validated short-form personality instruments to assess dark and light personality trait measures—the dark tetrad and a light "triad" of 3 personality dimensions were measured. Additionally, participants completed 3 tasks of interest that contribute to an understanding or one's ethics. Participants were administered the incentivized Social Value Orientation (SVO: Murphy et al., 2011) task to identify tendencies towards prosociality, and they were also administered the Coin Flip task (Houser et al., 2012) that presents a monetary temptation for dishonesty. Finally, participants made hypothetical choices across several scenarios of the classic Trolley Dilemma (Foot, 1967), where we also elicited self-reported mood ratings as a way to examine mood response differences to this dilemma across personality types. While the Trolley Dilemma assesses views regarding hypothetical moral dilemmas, scenarios are devised such that immoral acts of omission and commission can be identified in one's hypothetical choices, and these have used elsewhere to predict consequential anti-social choice (Dickinson and Masclet, 2019).

2. BACKGROUND

Dark personality traits, either individually or as a group, have been linked to various dimensions of unethical behavior, including fraud, deception, theft, bribe-taking, cyber-bullying, cheating, shoplifting, and marital infidelity (Nathanson et al., 2006a, 2006b; Zhao et al., 2016; Azizli et al., 2016; Sevi et al., 2020; Brown et al., 2019). The dark traits have been further linked to antisocial

tendencies or selfish behaviors in the domains of sex, power and money (Lee et al., 2013). While unethical behaviors have also been connected with certain dimensions of other personality structure models, such as the Big 5 (Tupes and Christal, 1992; Goldberg, 1993) and the HEXACO (Lee and Ashton, 2004), the dark tetrad of personality traits remains of interest and is easily integrated with a contrasting "light triad" of personality traits (Kantianism, Humanism, and Faith in humanity) that capture more positive dimensions of personality (Kaufman et al., 2019).

The dark side of personality links to unethical or morally questionable behavior due its connection to more primal behavioral tendencies. Dark traits have been found to correlate with spontaneous decision making (Čopková and Christenková, 2021), impulsivity and sensation-seeking (Crysel et al., 2013). Those with dark personality traits may also devalue collective interests and have a diminished concern for moral foundations (Jonason et al., 2015). As such, it is perhaps not surprising that researchers have found dark personality elements to be powerful antecedents to fraud behaviors (Harrison et al., 2018; Risenbilt and Commandeur, 2013), linked to overconfidence and risky corporate workplace behavior (O'Reilly and Hall, 2021; Olsen and Stekelberg, 2016), were positive predictors of antisocial online behavior (March and Marrington, 2021), and promote an increased likelihood of involvement in and ethical misconduct scandal (Van Scotter and Roglio, 2020). Such findings point to decision making in the moral domain that differs among those possessing dark personality characteristics.

Despite the strong intercorrelation between dark personality measures, the link between narcissism and immoral choice is less clear (see meta-analysis in Muris et al., 2017). For example, Zuo et al. (2016) found that narcissism may be *positively* associated with personal morality, at least among those with low self-esteem. Others have reported a link between Machiavellianism and psychopathy, but not narcissism, and low moral development and/or moral disengagement (Campbell et al., 2009; Egan et al., 2015). And, regarding cyberbulling, while some have connected all 4 dark traits to cyberbulling (Brown et al., 2019), others found that only psychopathy, sadism and Machiavellianism predicted this aversive behavior (Buckels et al., 2014). Narcissism is generally not viewed as poorly in society as psychopathy and Machiavellianism (Rauthman and Kolar, 2012), which somewhat distinguishes it from the other dark traits. To be clear, research has connected narcissism and unethical behavior (e.g., Van Scotter and Roglio, 2020; Rijsenbilt and Commandeur, 2013), but the link appears more

conditional at times and in ways that lead to narcissism as being perceived as a bit "less dark" perhaps.

There is more limited research on the light triad personality measures and moral choice. Light traits have been shown to negatively correlate with selfishness and aggression, and positively correlate with (socially beneficial) Dictator game donations (Kaufman et al., 2019). Others have found the light triad measure predicted decreased attitudes towards romantic relationship infidelity attitudes (Sevi et al., 2020) or increased likelihood of prosocial online behaviors (March and Marrington, 2021). Some have highlighted potentially beneficial qualities among those with dark personality traits, such as increased creativity (Kapoor and Kaufman, 2022), an increased willingness to make utilitarian decisions in moral dilemma (Karandikar et al., 2019), or a desire to be moral to preserve an inflated self-concept. When dark traits are linked with creativity, it may be malevolent rather than benevolent creativity that is encouraged. Also, the connection between dark personality and utilitarianism is likely more tenuous than it may appear from the literature.[1] And, any positive association between narcissism and personal morality appears conditional, at best (Zuo et al., 2016).

It is also worth noting that mood regulation is often connected to dark personality traits. Psychopathy, in particular, has been linked to poorer emotion regulation strategies (Walker et al., 2022), and emotion regulation difficulties are typically connected with pathological personality traits, in general (Pollock, 2016). Others have found the psychopathy and Machiavellianism, but not narcissism, predict moral disengagement when examining the dark triad (Egan et al., 2015). Though the research appears more limited regarding sadism and emotion dysregulation, it seems there may not be a connection here (Zeigler-Hill and Vonk, 2015). Given the importance of emotional engagement and mood in understanding moral judgments and choice (Greene et al., 2001), this present study is also interested in furthering our understanding of any differences in mood response to moral judgements among those with more light versus dark personality traits.

3. METHODS

[1] A classic moral dilemma, such as the Trolley, dilemma asks if one is willing to flip the switch to save five individuals but one will be killed as a direct result of the act of slipping the switch. Thus, the choice confounds Utilitarianism with those who simply prefer to be directly responsible for a death(s) rather than passively stand while others may die. Our particular design of Trolley dilemmas will help separate a true Utilitarian from someone who may flip a switch for more morally dubious reasons.

The study was preregistered prior to data collection on the Open Science Framework (hypotheses, design, variables, analysis).[2] The preregistration included a larger set of hypotheses than what are covered in this paper, as some of these additional hypotheses were not direction related to the questions of personality measures (i.e., dark or light measures) and choices on the ethical or moral dilemma tasks of interest here. The addition hypotheses and the associated data analysis and results of these can be found in the online Supplemental File. Our focus here is on the question of dark or light personality and moral choice as measured by decisions in two incentivized tasks and a classic (hypothetical) moral dilemma. Here, the key elements to the study methodology are described.

The study was administered on the Prolific platform (Palan and Schitter, 2018; Peer et al., 2017) with a target sample size of 2000-2500 observations. This target was set, in large part, based on the likely proportions of more extreme "dark" types that appear to be rare (Kaufman et al., 2019), and the desire to obtain at least n=100 participants who possess an extreme personality measure of this sort. Participants were recruited from the populations of US and UK residents to participant in a study asking about personality traits, mood, personal or political preferences. It was also noted that participants would administered 3 short decision tasks, 2 or which generated a bonus payment that would depend on participant decisions. The bonus payment was in addition to the flat payment received by all participants who completed the short study.

After measurement of some demographic characteristics and assessing baseline mood, the following key personality instruments were administered: short-form versions of the dark triad personality measures (subclinical psychopathy, narcissism, Machiavellianism: Jones and Paulhus, 2014), subclinical sadism (Plouffe et al., 2017), and the light triad personality measures (Kantianism, faith in humanity, humanism: Kaufman et al., 2019). The survey also administered the short-form 10-item version of the Big 5 personality inventory (the TIPI: Gosling et al., 2003), a 6-item cognitive reflection task to assess thinking style (Primi et al., 2016), and a visual measure of time discounting (Hershfield et al., 2012). Along with these, the survey administered the key decision tasks described next in more detail (SVO, Coin Flip task, and Trolley dilemmas). These tasks, along with the personality instruments, were randomized in order within the study. There was also a reassessment of mood measures following the Trolley dilemma choices that could be compared to the initial baseline mood measure given at the start

[2] The preregistration can be found at https://doi.org/10.17605/OSF.IO/A8QVD

of the study. In each instance mood was assessed along a 7-point scale across a set of positive (happy, enthusiastic, interested, determined, proud) and negative mood dimensions (irritated, distressed, ashamed, angry, sad).

3.1 Decision tasks

3.1.1 The Coin Flip task

The coin flip task we administered is a 10-flip version of the original task seen in the literature (Houser et al., 2012). Participants are asked to locate a coin and something to write with before progressing to the page that specifies exactly how outcomes are linked to payoffs. The main ask page then asks participants to flip the coin 10 times and report the number of HEADS flipped using a slider bar. The instructions further described that the payoff on the task, which contributes to their bonus payment for the study on Prolific, would be $0.15 for each HEADS reported. While seemingly small, the study in total took less than 15 minutes and the potential for an additional $1.50 payoff from the Coin Flip task would be more than their fixed pay for the short study. On the next survey page, participants were asked to input the specific outcome of each coin flip in order. This tasks therefore presents the participants with a monetary temptation to over-report the actual number of HEADS flipped (as well as possibly not flipping a coin at all, which can be assess through analysis of the response times on the task page). As has been noted in the literature, this task cannot identify cheaters at the individual level, but others have found that those who report more HEADS on this multi-flip Coin Flip task are also those who are more likely to have cheated in a separate task where the individual cheater is identifiable (Dickinson and Masclet, 2021).

3.1.2 The Social Value Orientation (SVO) task

Because ethical contexts often invoke the question of how others are affected, it is of interest to consider how different orientations towards prosociality are related to ethical decision making. Specifically, we administered the social value orientation (SVO) task (Murphy et al., 2011) to document where one's preferences lies along a spectrum that considers both individualist and pro-social orientations (with competitive and altruistic preferences at the extremes). The task elicits preferences across a series of allocations that describe one's own payoff as compared to another's payoff in the allocation. In fact, one may consider a test of the relationship between

one's SVO measure and outcomes in the ethical choice domain to be a test of the predictive validity of the SVO.[3]

3.1.3 The Trolley Dilemma task

We presented participants with a particular set of classic Trolley dilemma scenarios that would allow for the identification of more morally dubious behavior that can be done with the traditional classic Trolley dilemma (Foot, 1967). In the traditional Trolley dilemma, there 5 individuals who will be killed by a runaway train. Pulling a lever will divert the trolley to a side track and save these individuals but directly result in the death of another person on the side track. Thus, a confound in the Trolley dilemma not typically recognized is that a utilitarian choice (i.e., pull the lever to save a net 4 lives) is also the choice that makes one directly responsible for a death, rather than passively responsible. Because individuals tend to prefer *not* being directly responsible for someone's death, what seems to be the morally responsible choice to save the most lives is entangled in another moral challenge. While this is the very nature of the Trolley dilemma, it critically means that utilitarian choices are not so easily identified as choices made only for the greater social good.[4]

If the traditional dilemma saves 5 lives at the expense of 1 life, let us refer to this as a (5:1) Trolley dilemma. In our task, participants are asked to make the choice to pull the level or not for the following set of 6 Trolley dilemma scenarios: (5:1), (1:1), (5:0), (5:5), (2:1), (1:0). Figure 1 shows an example of the visual stimulus presented to the participant (in this case, the (1:0) dilemma. The usefulness of this particular set of Trolley dilemmas is that several choice types can be distinguished from one's responses. Trolley scenarios (5:1), (5:0), (2:1), and (1:0) are all dilemmas where it would be utilitarian to pull the lever. To *not* pull the level in scenario (5:0) and/or (1:0) would be particularly troubling, however. Here, there is no confound between the utilitarian choice and the possibility that one would prefer direct responsibility over the death of another. Failure to choose "pull the lever" in scenario (5:0) or (1:0) can be considered an immoral act of omission. Additionally, there is no benefit in terms of lives saved/lost in choosing the pull the lever in Trolley scenarios (5:5) or (1:1). Since it has been found that most

[3] Murphy et al. (2011) document evidence of predictive validity in terms of SVO and cooperation in Prisoner's Dilemma games.

[4] Additional versions of the trolley dilemma may further alter the environment by make it more personal (i.e., "push" an individual rather than pull a lever, or identify the types of people who may live or die).

prefer to *not* be directly responsible for one's death (versus passively allowing it, see Descioli et al., 2011, on this point), we can consider it an immoral act of commission to pull the lever in either of these two Trolley dilemma scenarios. Thus, our particular set of Trolley dilemmas allows for a more precise and useful categorization of one's ethics in this particular hypothetical choice task.

4. HYPOTHESES

Preregistered hypotheses were developed from the existing literature (discussed above) linking dark personality traits, particularly sadism, psychopathy, and Machiavellianism, to various forms of unethical or immoral conduct. Dark personality traits have been previously linked to less prosocial behaviors in at least one incentivized decision task. The Background section above also noted the differences in emotion regulation most consistently found in those with Machiavellian or psychopathic traits, and may contribute to reported utilitarian choices in moral dilemmas. Finally, the particular link between sadism and pleasure from another's pain, we hypothesize a positive relationship between sadism and immoral acts of omission and commission in the Trolley dilemma. The following hypotheses related to this study were preregistered on the Open Science Framework prior to data collection.[5]

Hypothesis 1: Those with higher levels of dark personality traits will have lower levels of pro-sociality, as measured by the SVO (social value orientation measure).

Hypothesis 2: HEADS reported on the coin flip task will be positively (negatively) related to dark (light) personality traits.

Hypothesis 3: Dark personality traits, in particular Machiavellianism and psychopathy, will predict utilitarian choices.

Hypothesis 4: Higher levels of psychopathy and Machiavellianism will predict reduced levels of baseline and post-Trolley task self-reported emotion, as well as less emotion-level change.

Hypothesis 5: Sadism will predict a greater likelihood of an immoral act of omission or commission (in the Trolley dilemma).

[5] The full set of hypotheses, which includes others not directly related to this paper, can be found in the preregistration document at https://doi.org/10.17605/OSF.IO/A8QVD. The hypotheses are listed here may be ordered differently than in the preregistration document for exposition purposes. Preregistered hypotheses not directly related to this paper are examined in the Online Supplement Information Appendix.

Hypothesis 6: committing immoral acts (choices) of omission or commission in the Trolley dilemma will predict the # HEADS reported in the Coin Flip task.

4.1. A simple framework for moral decision-making

While the testable hypotheses were generally derived from previous empirical findings, a general framework for decision making in the moral domain is useful as a way to help organize our understanding of underlying mechanisms. Researchers have noted the importance of social or personal norms in understanding choice (see, e.g., Bicchieri, 2005; Elster, 1989; Dubreuil and Grégoire, 2013), but these frameworks typical do not consider that personal norms may be considered immoral or unethical, in general. One example in the literature has suggested that dark personality types differ in subjective norms as a reason for unethical intentions (in this case, contract cheating: see Curtis et al., 2022). While there are a variety of available frameworks that penalize utility for deviations from some ideal behavior, consider the following utility function (Masclet and Dickinson, 2019).

$$U(a) = b(a) - c(a) - v(a - \hat{a})$$

Here, utility for a particular behavior or action, a, is not only a function of benefits, $b(a)$, and costs $c(a)$, but there is also a disutility association with deviations from some target behavior, $\hat{a}$, via the term $v(a - \hat{a})$. A general specification such as this allows one to have any target behavior within her preferences, moral or immoral. The specification can be modified and one's target action may or may not be subject to social influences (see Masclet and Dickinson, 2019), but this framework is simple and can describe many of our hypotheses as being a result of hypothesizing that dark personality types have different morals or personal norms reflected in $\hat{a}$.

5. RESULTS

In total, n=2565 participants completed the study via the Prolific platform (n=1261 residing in the UK, n=1304 residing in the US). Of these, n=2463 participants (n=1203 UK, n=1260 US) passed the attention checks within the survey. Age and sex were not elicited on the survey but were downloaded from the available characteristics provided on all study participants from Prolific. Some participants did not have such data available (participants are allowed to withdraw consent to share those data with researchers), and so the sample available with age and sex was n=2413 (n=1178 UK, n=1235 US). We preregistered plans to analyze data from those

who had no missing survey data, but we did not anticipate that some profile data may be incomplete from Prolific. In general, we present all results from estimations that do and do not include demographic characteristics as control variables as a robustness check.

Figure 2 shows the general tendencies in our sample regarding average measures of light and dark personality measures. Each dimension of the dark tetrad or light triad measures on a similar 1-5 scale, and so scores for each dimension were average across the light triad and dark tetrad measures. Differencing these two can be used to generate a measure of *relative* lightness or darkness of personality along these dimensions. These data in Figure 2 are similar to those reported in Kaufman et al., (2019) in that they show tendencies towards a *relatively* more light over dark personality type. Some nontrivial level of dark personality traits is not uncommon, even among those who are score higher on the light triad of personality measures. Though preregistered analysis of hypotheses will be conducted on the full sample, some exploratory analysis will focus on participants in the approximate 10% tails of our sample in terms of *relatively* most light or dark, which are highlighted in the shaded regions in Figure 2.

Hypothesis 1 test—Dark personality and prosociality

To examine the impact of dark personality traits on prosociality, the dependent variable is the continuous measure one's "*SVO angle"* as derived from the SVO task and as calculated in Murphy et al. (2011). Here, the higher one's *SVO angle*, the higher one's orientation towards prosociality. We estimated models that linear regression models with the *Dark Triad* or *Dark Tetrad* as key measure of dark personality tendencies. Simple regressions are compared next to regressions that control for age and sex, and that additionally control for Big 5 personality traits. Finally, though we did not preregister hypotheses relating light personality traits to prosociality, in the spirit of Hypothesis 1, we also estimated models that regressed *SVO Angle* on the *Light Triad* measure, as well as on the *NetLight* measure (i.e., *NetLight = Light Triad – Dark Tetrad*). The results are shown in Table 1, and they support Hypothesis 1.

Finally, in order to examine the individual dark versus light personality trait measures' impact on prosociality, we regressed similar models to those in Table 1 with the key independent variable being the specific dark or light personality trait. These models include the full set of control variables, and we summarize the results by means of the coefficient plots shown in Figure 3 (see online Appendix A for full estimation results). From these additional estimations,

we see that each dark trait predicts a significantly lower *SVO Angle*, and each light trait predicts a significantly higher *SVO Angle.* In short, the data are consistent with the specific preregistered Hypothesis 1, but more generally they also show that light personality trait,s as well as the relative lightness over darkness of one's personality, are associated with a more prosocial personality orientation.[6]

Hypothesis 2 test—Coin Flip task and dark/light personality

Models similar to those used to evaluate Hypothesis 1 were used to examine Hypothesis 2, with the difference being the dependent variable, *HEADS*. The results are reported in Table 2 and show that the dark personality clusters are stronger predictors of higher reported *HEADS* outcomes in the Coin Flip task than light personality trait clusters. Given the lack of significant predicted coefficient on the *Light Triad* coefficient estimate in model (5), we also conclude that the significant of the *Net Light* coefficient estimate in model (6) is a result of the *HEADS* report variance predicted by dark personality trait clusters. Because dark personality traits predicted increased *HEADS* reported but light personality traits did not predict decreased *HEADS* reported, we conclude there is partial support for Hypothesis 2.

Figure 4 (see also Appendix Table A2) shows coefficient plot results from estimation results for regressions of *HEADS* reports on the separate individual personality trait measures. At the individual trait level, the dark personality traits have a more marginal significance in predicting higher reported *HEADS* as compared to when the traits were considered in clusters (i.e., the dark triad or tetrad).

***Hypothesis 3 test*—Dark personality traits, in particular Machiavellianism and psychopathy, will predict utilitarian choices.**

Table 3 shows results of models regressing the proportion of utilitarian choices made in Trolley dilemma scenarios (1:0), (2:1), (5:0), and (5:1). We therefore score *Proportion Utilitarian Choices* equal to .0, .25, .50, .75, or 1.00. We excluded Trolley dilemma (5:5) and (1:1) from this analysis give the lack of a unique utilitarian choice. Though previous studies have reported data to indicate that dark personality traits may favor utilitarian choices in the Trolley dilemma

[6] Results are similar if restricting the analysis to those who are highest and lowest on the *NetLight* scale (i.e., the shaded region of participants in Figure 2. Results are available on request.

(e.g., due to the emotion detachment that may help one make a difficult choice in such a dilemma), our data fail to support Hypothesis 3. In fact, it appears the light personality traits are found to predict utilitarian choices in our data, while dark personality traits do not predict the proportion of utilitarian choices. Because the set of 4 Trolley dilemmas analyzed included dilemmas that suffer the criticism of confounding utilitarianism with direct responsibility for a death (i.e., the typical (5:1) and (2:1) dilemmas), we also conducted similar analysis using just the (1:0) and (5:0) dilemmas that would most unambiguously identify utilitarian choice and our results were similar (these results are available on request). Separate regressions on specific individual dark or light personality traits similarly find no predictable relationship between the dark personality traits and utilitarian choice, while it is specifically the traits of "humanism" and "faith in humanity" that predict a higher proportion of utilitarian choices made (see online Appendix Table A3).

Hypothesis 4 test: **Psychopathy and Machiavellianism will predict lower baseline mood and post-Trolley mood levels, as well as a lesser mood level change pre- to post-Trolley dilemma.**

The hypothesis preregistered focused on the specific dark traits of psychopathy and Machiavellianism, though we present estimation results that examine all 4 darks traits and their association with self-reported mood. Table 4 shown results from estimations where an average of baseline mood is regressed on dark personality traits in succession, with controls for the Big-5 traits, age, and sex. The mood variable is the average positive mood (averaged 7-point Likert-scale self-ratings of the following emotions: happy, enthusiastic, interested, determined, proud) or average negative mood (averaged 7-point Likert-scale self-ratings of the following emotions: irritated, distressed, ashamed, angry, sad) assessed at the beginning of the survey (prior to any decision task administration). Each of the dark traits predicts an increased level of baseline negative mood, while Narcissism also predicts increased levels of baseline positive mood. Female and younger participant is associated with lesser baseline levels of positive mood, whle younger participants also predict higher baseline negative mood. In general, aside from the Big-

5 trait of *openness*, all other Big-5 traits predict higher positive and lower negative baseline mood. Results in Table 5 show similar estimations with the dependent mood variables being those elicited immediately after completing the Trolley dilemma task.

Because some variables are found to predict an increase in positive or negative mood both before and after the Trolley task, we also constructed a variable measuring the change in one's net *positive-negative* mood reports from baseline to post-Trolley task. This variable, *Net Positive mood change*, is used as the dependent variable in a series of similar regressions, where we also considered the light-triad personal traits as regressors. These results are shown in full in the online Appendix Table A4, but we summarize the key personality trait effects also in Figure 5 (which also summarize the key estimated dark trait effects from Tables 4 and 5). Overall, our data fail to support Hypothesis 4, which stated that psychopathy and Machiavellianism would predict lower baseline and post-Trolley mood, as well as a lesser mood change after completing the Trolley task. However, the results suggest mood dysregulation is perhaps the more potent variable to consider. The significant estimated mood findings indicate that psychopathy and Machiavellianism stand out in that they predict a unique (and, perhaps disturbing) mood shift toward increased positive and reduced negative mood after going through the hypothetical life-and-death moral dilemma task.[7]

Hypothesis 5 test: **Sadism will predict a greater likelihood of an immoral act of omission or commission (in the Trolley dilemma).**

The key outcome of what we defined as immorality in the Trolley dilemma is that individuals would be either unnecessarily killed (immoral act of omission) or killed preferentially by one's

[7] These results are similar if using the Dark or Light cluster measures in place of individual trait measures to assess the impact of personality trait. That is, higher scores on the Dark Tetrad predict an increase in one's net positive mood change, and higher scores on the Light Triad predicts a reduction in one's net positive mood after completing the Trolley task (as does one's *NetLight* personality score). In short, more light personality types have a more negative change in their self-reported mood after completing this task, while mood turns more positive after this task for those with more dark personalities (and those effects seem driven largely by psychopathy and Machiavellianism scores).

action rather than inaction (immoral act of commission). As such, these acts of (hypothetical) immorality seem most linked to the trait of sadism among the dark traits. Hypothesis 5 is tested in a serious of linear probability regressions that defined the dependent variable as an immorality indicator variable.[8] That is, we set *Immoral omission* =1 if the participant chose the immoral action to not "pull the lever" in either or both of the Trolley dilemmas (5:0) and (1:0)). *Immoral commission* was set equal to 1 if the participant chose the immoral action to actively pull the lever in either or both of the Trolley dilemmas (5:5) and (1:1). We show in Figure 6 the coefficient plots summary of the key estimated effects of sadism on the likelihood of committing an immoral act. Models were run with and without control variables for sex, age, and the Big-5 personality traits, and we also include comparison results that use the *NetLight* personality measure in place of the *sadism* personality control. The results support Hypothesis 5 in that those with high levels of the sadism trait are significantly more likely to commit an immoral act of omission and commission in the Trolley dilemma.

Though the preregistered hypothesis specified sadism as the key personality trait of interest, we also estimated similar models examining each of the individual dark or light personality trait effects on Trolley immorality. Each of the 4 dark personality traits positively predicted the likelihood of choosing to pull the lever in the Trolley (1:1) and/or (5:5) dilemmas (i.e., immoral act of commission), while light personality traits were statistically insignificant in these estimations. Regarding choices to *not* pull the level in Trolley (1:0) and/or (5:0) dilemma, psychopathy and sadism predicted an increased likelihood, while humanism and faith in humanity predicted a reduced likelihood, of choosing this immoral act of omission. These results are reported in full in the online Appendix Tables A5 and A6.

Hypothesis 6 test: **Immoral acts (choices) of omission or commission in the Trolley dilemma will predict the # HEADS reported in the Coin Flip task.**

The essence of Hypothesis 6 is to test whether presumed immorality in the hypothetical choice dilemma predicts the likelihood of dishonesty in the consequential task. For these regressions, the indicator variable for an immoral act of omission or commission is used as an independent variable to predict the dependent variables *HEADS* reported in the Coin Flip task. Figure 7

[8] Virtually identical results in sign and significance are found using nonlinear Probit estimations, and so we report linear probability regression results for simplicity and ease of interpretation.

shows the summary of these estimations results for models that varied in the set of control variables. The data fail to support Hypothesis 6 as we find no statistically significant impact of one's Trolley immorality choice on predicting *HEADS* reported. The full model results are shown in the online Appendix Tables A7, which highlight that main Coin Flip task results in Table 2 that younger participants and those with higher measures of *NetLight* predict higher *HEADS* reports.

6. DISCUSSION

The data supported some of our preregistered hypotheses (H1, H2, and H5), failed to support others (H3 and H6), and revealed somewhat related support for H4 in the sense of highlighting that dark personality traits may differ in their mood response to the Trolley dilemma task. The overall theme of our results suggests that dark personality traits are related to lower prosocial tendencies and an increased likelihood of making choices likely considered immoral in both hypothetical and consequential decision environments. While others have reported an increase in utilitarian choices by individuals possessing dark personality traits (e.g., Bartels and Pizarro, 2011), our results do not support this hypothesis. To our knowledge all previous studies reporting this relationship between dark traits and utilitarianism do so using Trolley or similar moral dilemma environments where there is a confound between the utilitarian choice and direct responsibility for others' deaths. We removed this confound in a subset of our Trolley dilemmas, but even if we restrict our analysis to dilemma that contain this confound, our results do not support previous findings regarding dark personality traits and Utilitarian choice. Rather our data show robust evidence that those with more light personality traits are more likely to make the utilitarian choice. This is true even though the more light personality types experience a net-worsening of their mood after completing the Trolley dilemma task

How mood relates to dark personality traits in these hypothetical moral dilemmas is curious. Our specific preregistered mood hypothesis was not supported, but exploratory analysis identified a stark difference between the mood shift reported during before and after the administration of the Trolley dilemma task for those who were more light versus dark in their personality traits. Whether this mood shift could explain choices is an interesting one. Exploratory analysis did not find a significant direct impact of one's "Trolley mood shift" on the likelihood of *Immoral omission* or *Immoral commission* choices. However, instrumental

variables analysis showed increases in post-Trolley positive mood explained by *lower* values of *NetLight* (i.e., higher values of *NetDark*) significantly predicted a greater likelihood of *Immoral Omission* choices ($p = .02$), *Immoral Commission* choices ($p < .01$), and a greater number of *HEADS* ($p < .01$) (see Appendix Table A8 for full results). Though our main hypothesis test of H6 did not support the direct connection between hypothetical Trolley dilemma choices and consequential choices in the Coin Flip task, these exploratory results suggest further research is needed to identify the extent to which extent there is a link between mood response, hypothetical moral dilemma choices, and consequential choices in the ethical domain. This is important given past research reporting that choices in hypothetical dilemmas do not predict in (somewhat) similar consequential versions of the dilemma (Bosyn et al., 2018), but yet others have found they may be predictive of antisocial choice in other consequential environments (Dickinson and Masclet, 2019).

Of course, important limitations of this study should be recognized. If self-report measures are to be used to evaluate personality for such purposes as making judgments regarding one's ethics, then participants may wish to respond in socially acceptable or desirable ways bias (Zerbe and Paulhus, 1987). One may respond to hypothetical dilemma questions or make choices in consequential temptation tasks in a way that is deemed to align more with acceptable norms. While we did not take into account or attempt to measure any social desirability bias (e.g., see Stöber, 2001), the anonymous nature of the online study administered through a 3rd party platform likely limits the concern of such social desirability bias in the data. To the extent that it may exist, its presence would imply that our evidence of dishonest or immoral choice tendencies are likely underestimated. If dark personality types are less impacted than more light personality types by the desire to submit responses deemed socially acceptable, but no more likely to actually make dishonest choices, then our results could overestimate the extent to which dark traits are actually responsible for more dishonest behaviors.

This study did not set out to compare ethical behaviors across cultures, and so no hypotheses were preregistered to compare results in the US and UK subsamples of our data. However, if one conducts our analysis on the separately subsamples of participants from each location (what a robust sample size of over n=1200 per location), then we find that some notable differences are estimated. The significant results found to support H1, H2, and H5 are due to the US subsample of participants. However, the surprising exploratory finding that showed an

increase in positive mood after completing the Trolley task was only statistically significant in the UK subsample of participants. These findings may suggest an interesting influence of one's culture or societal norms on the link between personality and behavior. While this was not the focus of the present study, it would seem to merit more systematic investigation going forward.

The cross-sectional nature of the study implies our results are more indicative of correlations or associations, as opposed to causation. This is a criticism common to much of the research in this area, as any proper experimental design would seek to randomly assign, or manipulate in a within-subjects manner, the key predictor of interest: personality type. We therefore hope that our study will be an important contribution to a literature that can identify systematic relationships between personality and moral choice using many complementary approaches.

Finally, the present paper only explored a limited set of decision environments that do not capture the full breadth of how unethical behaviors may manifest, and they do not examine ethics in less anonymous situations. For example, those influenced by one's immoral choices may be more or less socially distant with the decision making, or the choice itself may be more or less observable by others. Researchers have articulated a difference between social norms versus moral (personal) norms (see Bicchieri, 2005; Elster, 1989; Dubreuil and Grégoire, 2013), and the present research did not seek to identify or measures participants' perceptions of such norms. Are differences in ethical choice by personality type due to key differences in views of social or moral norms, or are they due to differences in concern for such norms. Some recent research has suggested that dark personality types differ in subjective norms as a reason for unethical intentions (in this case, contract cheating: see Curtis et al., (2022). This would seem another fruitful area where additional research is needed. While these (and likely other) limitations point to the need for additional work, the present paper hopes to have contributed to our understanding of how dark versus light personality traits may be associated with socially undesirable behaviors or choices in the face of moral dilemmas.

FIGURE 1: Trolley Dilemma visual stimulus used in survey
(1:0) dilemma—pulling lever saves 1 life on main track and no one is killed on the side track

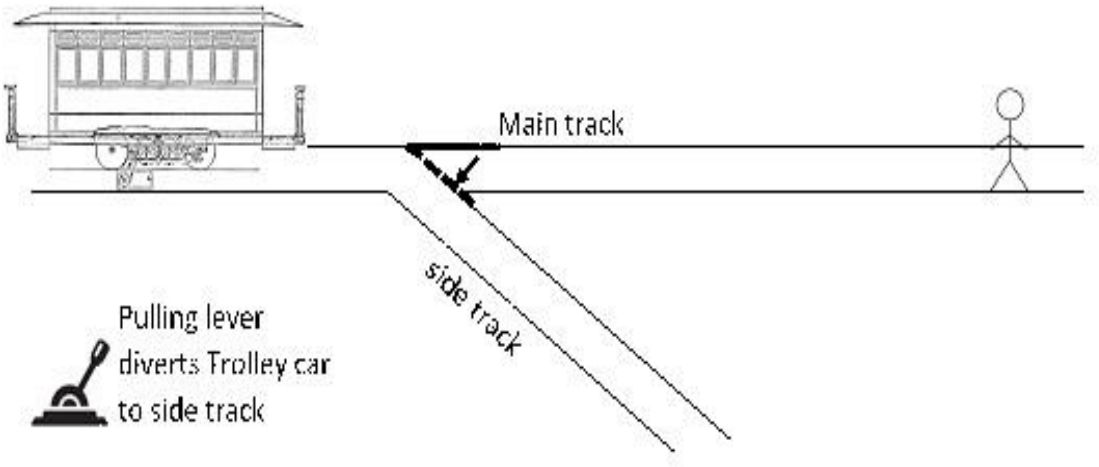
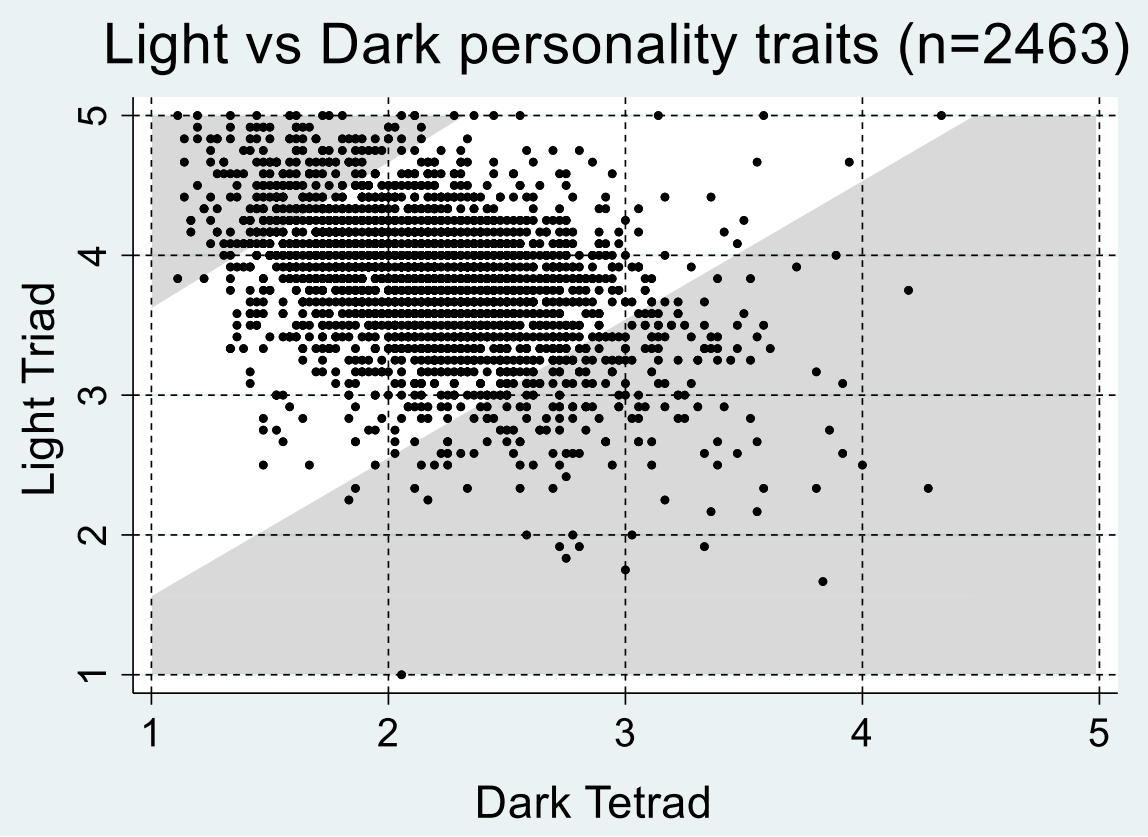
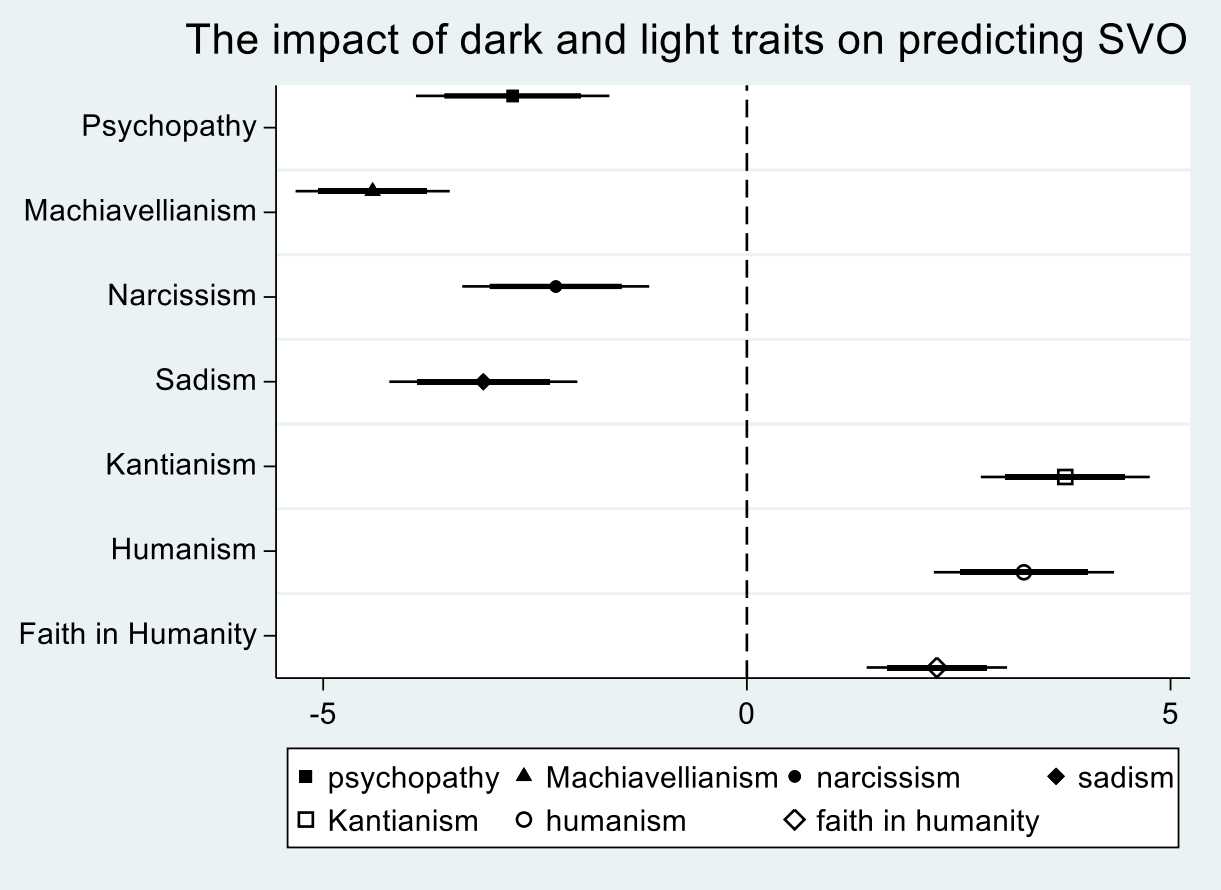


FIGURE 2: Personality Traits in study sample



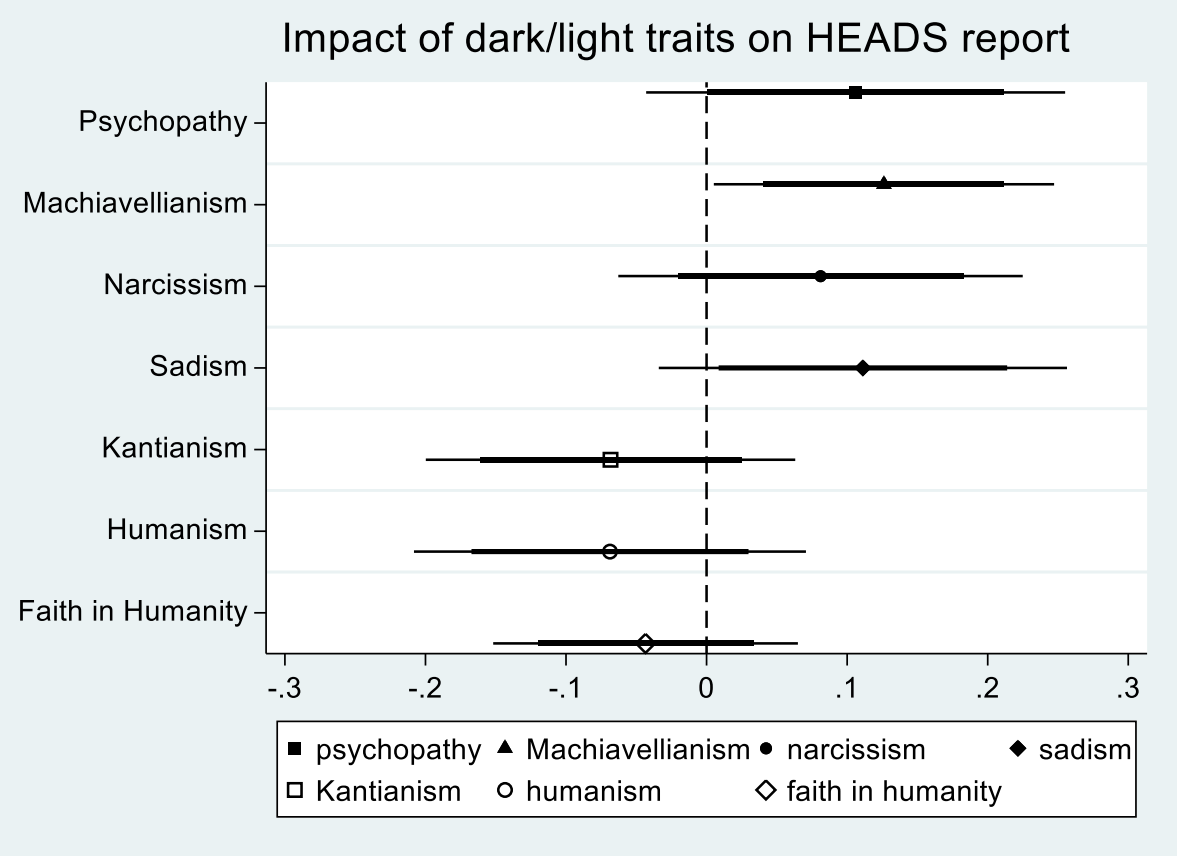
Notes: Shaded regions show the approximate cutoff for the upper and lower 10% of participants in terms of their Light-Dark average traits (NetLight > 2.65, n=234: NetLight < .55, n=245)

FIGURE 3: H1 coefficient plots—personality traits and SVO.



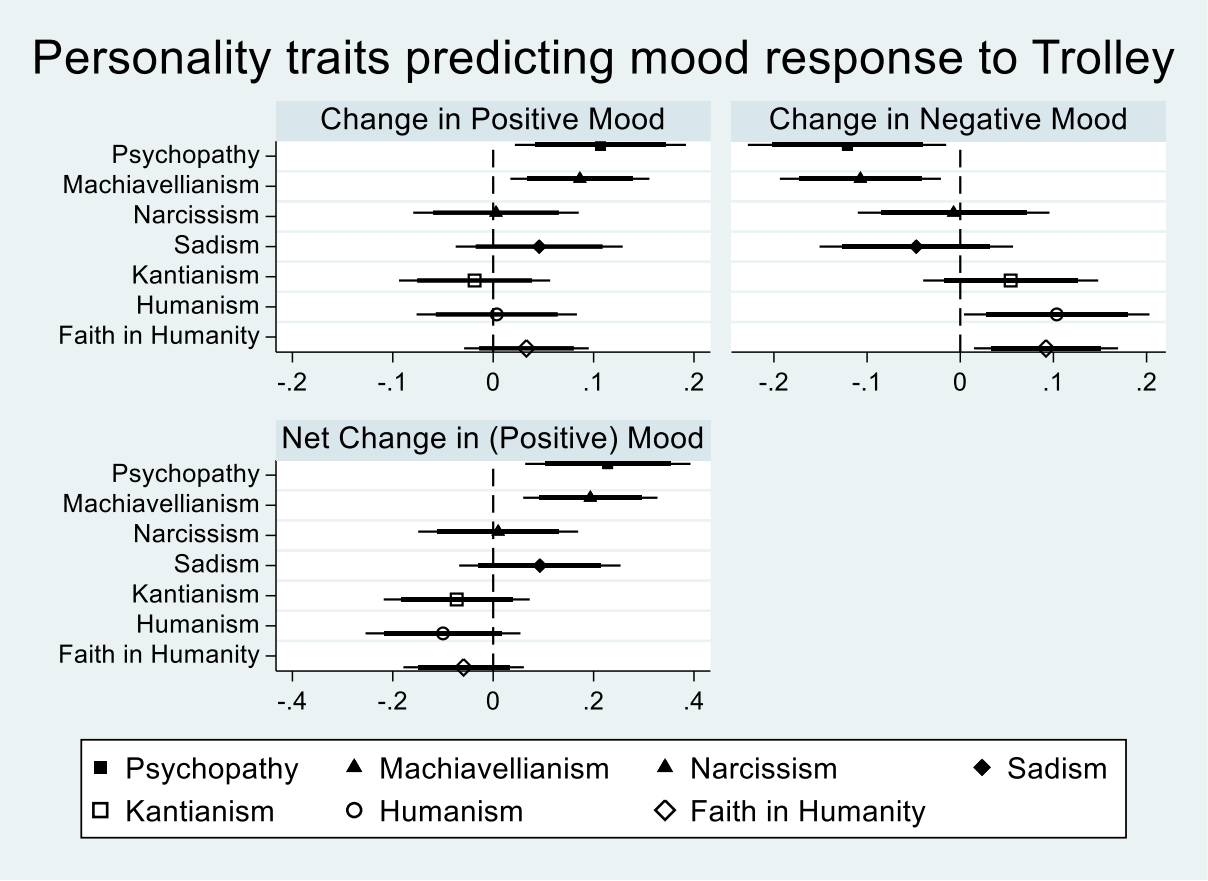
Notes: Bars show the 99% (thin bars) and 95% (thicker bars) confidence interval for a 1-tailed test (preregistered hypotheses) on the coefficient estimate of the personality trait's effect on *SVO Angle*. Models estimated include controls for age, sex, and the Big 5 traits.

FIGURE 4: H2 coefficient plots—personality traits and Coin Flip task.



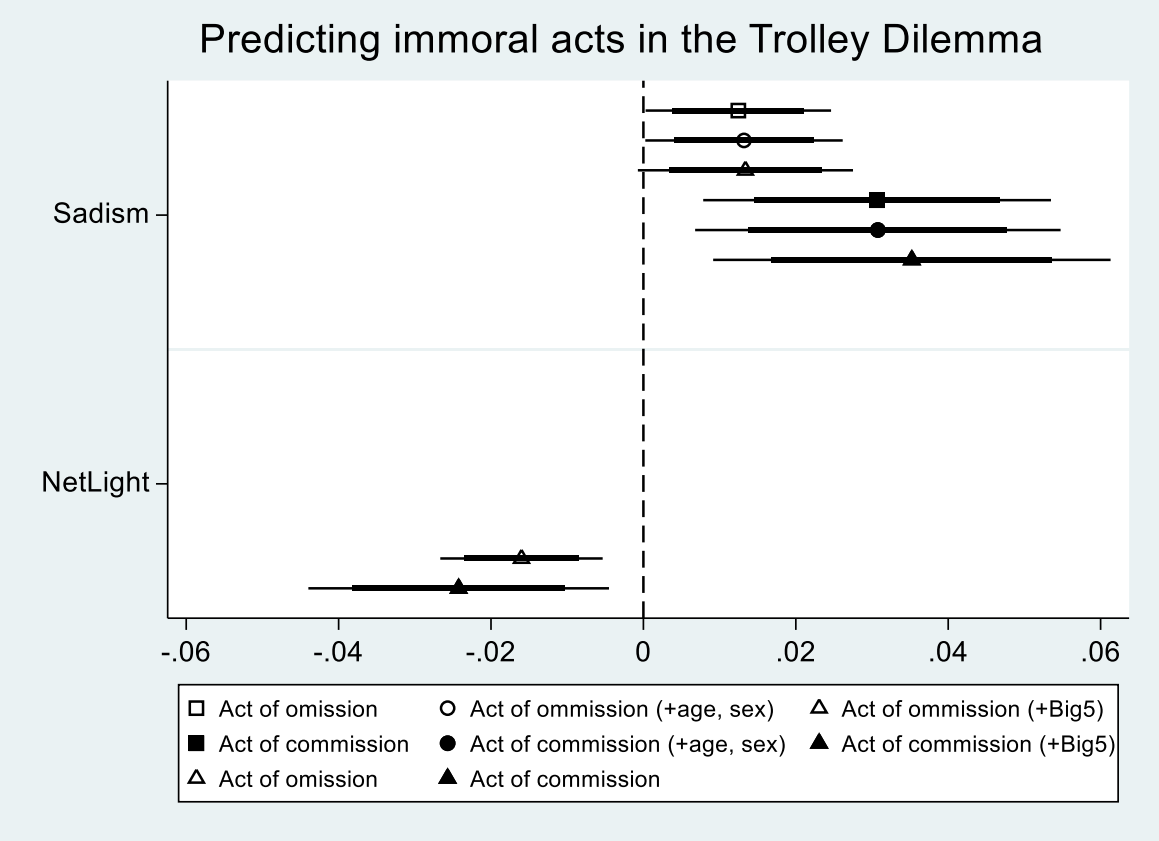
Notes: Bars show the 99% (thin bars) and 95% (thicker bars) confidence interval for a 1-tailed test (preregistered hypotheses) on the coefficient estimate of the personality trait's effect on *HEADS* reported in the Coin Flip task. Models estimated include controls for age, sex, and the Big 5 traits.

FIGURE 5: Differential mood response to completing the Trolley dilemma



Notes: Bars show the 99% (thin bars) and 95% (thicker bars) confidence interval for the 2-tailed test on the coefficient estimate of the personality trait's effect on one's self-reported mood change from baseline to immediately after the Trolley dilemma task. Two-tailed tests confidence intervals are shown in this instance given the results fail to support Hypothesis 4, and so we can assess the opposite-hypothesized result's significance at standard levels. Models estimated include controls for age, sex, and the Big 5 traits. Full results for the dark traits effects for the two upper panels are those in Tables 4 and 5. See Appendix table A4 for full estimation results of the Net Change in (Positive) Mood in the bottom panel.

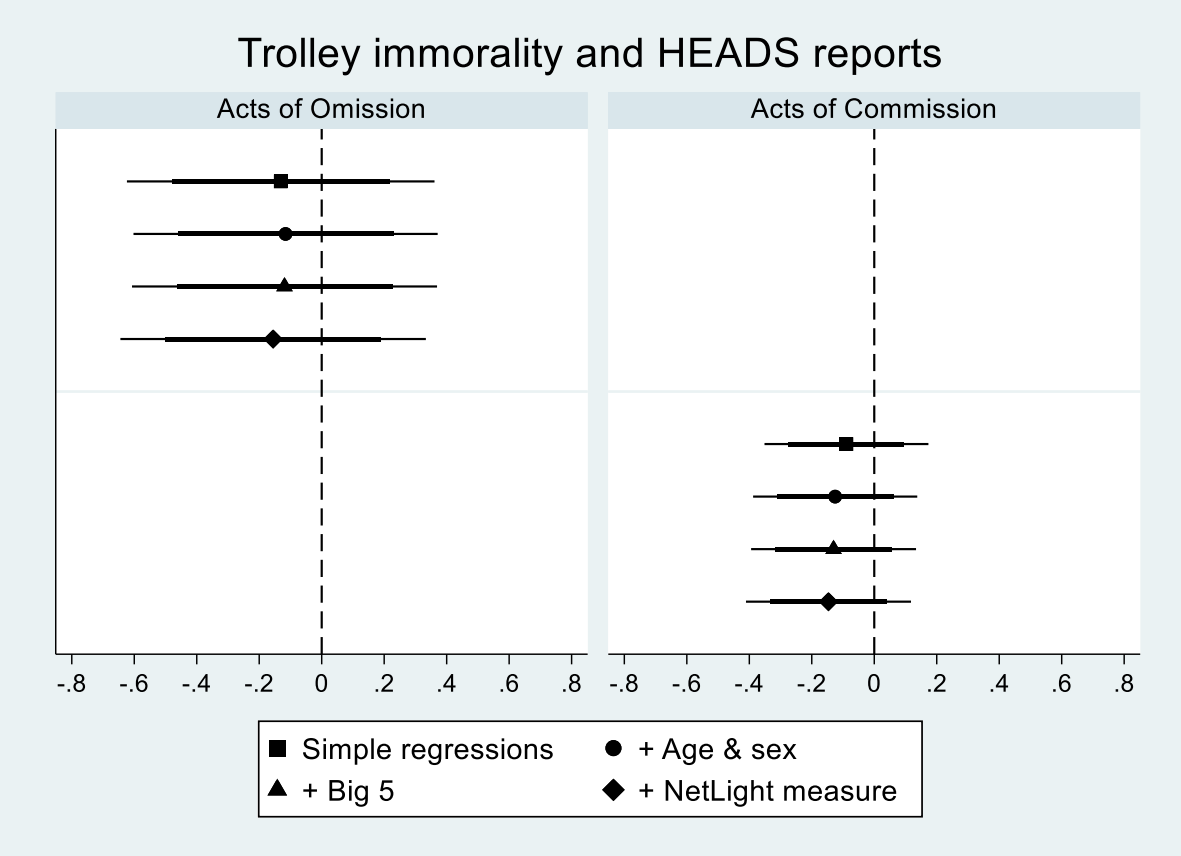
FIGURE 6: H5--Sadism (dark personality) effects on Trolley immorality



Notes: Bars show the 99% (thin bars) and 95% (thicker bars) confidence interval for the preregistered 1-tailed test on the coefficient estimate of the personality trait's effect on the likelihood of choosing an immoral act of omission or commission. See Appendix Tables A5 and A6 for analysis using individual-specific personality traits as regressors.

FIGURE 7: H6 test—Trolley immorality as predictor of HEADS reports (Coin Flip task)

(Dep variable=HEADS: Independent variable is indicator for those who made an immoral act of omission or commission choice in the Trolley dilemma task)



Notes: Bars show the 99% (thin bars) and 95% (thicker bars) confidence interval for the preregistered 2-tailed test on the coefficient estimate (2-tailed CI shown given point estimates are opposite the preregistered effect). See Appendix Table A7 for full estimation results.

TABLE 1: Hypothesis 1 tests—Dark traits and *SVO Angle*

Dependent Variable = *SVO Angle*	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Independent Variable	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**
constant	36.20 (1.23)**	35.15 (1.68)**	31.05 (2.40)**	32.07 (2.46)**	-28 (2.14)	16.29 (1.62)**
Dark Triad	-5.99 (.50)**	-5.61 (.53)**	-5.59 (.58)**	---	---	---
Dark Tetrad	---	---	---	-5.94 (.61)**	---	---
Light Triad	---	---	---	--	5.64 (.56)**	---
NetLight	---	---	---	--	---	4.38 (.35)**
Age	---	.003 (.02)	.008 (.02)	.004 (.02)	.046 (.02)*	.022 (.02)
Female (=1)	---	1.45 (.55)**	.66 (.57)	.58 (.57)	1.45 (.56)**	.68 (.56)
Extraversion	---	---	.54 (.17)**	.50 (.17)**	-.04 (.16)	.26 (.16)
Agreeable	---	---	.86 (.24)**	.75 (.24)**	.54 (.25)*	.15 (.25)
Conscientious	---	---	-.36 (.21)	-.42 (.21)*	-.36 (.21)	-.48 (.21)*
Emotional Stability	---	---	-.57 (.19)**	-.58 (.19)**	-.65 (.19)**	-.59 (.19)**
Openness	---	---	.31 (.21)	.25 (.21)	-.02 (.21)	.10 (.21)
Adjusted R-squared	.0543	.0547	.0667	.0673	.0699	.0884
Observations	2,463	2,413	2,413	2,413	2,413	2,413

Notes: *$p < .05$, **$p < .01$ for 1-tailed test of a preregistered directional hypothesis. Otherwise, 2-tailed tests *p*-values reported.

TABLE 2: Hypothesis 2 tests—Dark/Light traits and the Coin Flip task

Dependent Variable = *HEADS*	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Independent Variable	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**
constant	4.88 (.16)**	5.63 (.22)**	5.63 (.32)**	5.5 (.33)**	6.46 (.28)**	6.13 (.22)**
Dark Triad	.28 (.07)**	.18 (.07)**	.18 (.08)**	---	---	---
Dark Tetrad	---	---	---	.20 (.08)**	---	---
Light Triad	---	---	---	--	-.11 (.07)	---
NetLight	---	---	---	---	---	-.11 (.05)**
Age	---	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**
Female (=1)	---	-.10 (.07)	-.10 (.08)	-.10 (.08)	-.13 (.07)	-.11 (.07)
Extraversion	---	---	-.0004 (.02)	.001 (.02)	.01 (.02)	.01 (.02)
Agreeable	---	---	.01 (.03)	.01 (.03)	.01 (.03)	.02 (.03)
Conscientious	---	---	-.01 (.03)	-.01 (.03)	-.02 (.03)	-.01 (.03)
Emotional Stability	---	---	-.01 (.03)	-.01 (.03)	-.004 (.03)	-.01 (.03)
Openness	---	---	.01 (.03)	.01 (.03)	.02 (.03)	.02 (.03)
Adjusted R-squared	.0068	.0157	.0139	.0140	.0125	.0139
Observations	2,463	2,413	2,413	2,413	2,413	2,413

Notes: *$p < .05$, **$p < .01$ for 1-tailed test of a preregistered directional hypothesis. Otherwise, 2-tailed tests *p*-values reported.

TABLE 3: Hypothesis 3 tests—Dark/Light traits and Utilitarian Trolley dilemma choices

Dependent Variable = *Proportion Utilitarian Choices*	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Independent Variable	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**
constant	.93 (.02)**	.94 (.03)**	.92 (.04)**	.92 (.04)**	.81 (.03)**	.91 (.03)**
Dark Triad	-.01 (.01)	-.01 (.01)	-.007 (.009)	---	---	---
Dark Tetrad	---	---	---	-.008 (.010)	---	---
Light Triad	---	---	---	---	.03 (.009)**	---
NetLight	---	---	---	---	---	.02 (.006)**
Age	---	-.0001 (.0003)	-.0001 (.0003)	-.0001 (.0002)	.0001 (.0003)	-.0001 (.0003)
Female (=1)	---	-.006 (.009)	-.01 (.01)	-.01 (.01)	-.01 (.01)	-.01 (.01)
Extraversion	---	---	.003 (.003)	.003 (.003)	.01 (.003)	.003 (.003)
Agreeable	---	---	.008 (.004)*	.008 (.004)*	.003 (.004)	.003 (.004)
Conscientious	---	---	-.0004 (.003)	-.001 (.003)	-.001 (.003)	-.001 (.003)
Emotional Stability	---	---	-.006 (.004)*	-.006 (.003)*	-.006 (.003)*	-.006 (.003)
Openness	---	---	-.001 (.003)	-.001 (.003)	-.002 (.003)	-.001 (.003)
Adjusted R-squared	.0001	-.0005	.0006	.0006	.0064	.0038
Observations	2,463	2,413	2,413	2,413	2,413	2,413

Notes: *$p < .05$, **$p < .01$ for 1-tailed test of a preregistered directional hypothesis. Otherwise, 2-tailed tests *p*-values reported. The *Proportion Utilitarian Choices* variables takes on the value of 0, .25, .5, .75, or 1 and represents the proportion of utilitarian choices made across the 4 Trolley scenarios with a unique utilitarian choice (i.e., (1:1) and (5:5) Trolley scenario choices not considered).

TABLE 4: Hypothesis 4 tests—Psychopathy and Machiavellianism predictions on baseline mood

	Dependent Variable = *Positive Mood (baseline)*				**Dependent Variable = *Negative Mood (baseline)***			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Independent Variable	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**
constant	1.34 (.20)**	1.26 (.20)**	1.02 (.16)**	1.12 (.19)**	3.20 (.20)**	3.65 (.20)**	3.80 (.16)**	3.50 (.19)**
Psychopathy	-.02 (.04)	---	---	---	.29 (.04)**	---	---	---
Machiav.	---	.002 (.03)	---	---	---	.15 (.03)**	---	---
Narcissism	---	---	.14 (.04)**	---	---	---	.20 (.04)**	---
Sadism	---	---	---	.05 (.04)	---	---	---	.21 (.04)**
Age	.01 (.002)**	.01 (.002)**	.01 (.002)**	.01 (.002)**	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**
Female (=1)	-.18 (.05)**	-.18 (.05)**	-.16 (.05)**	-.17 (.05)**	.001 (.05)	-.06 (.05)	-.05 (.05)	-.04 (.05)
Extraversion	.18 (.01)**	.18 (.01)**	.15 (.02)**	.17 (.01)**	-.03 (.01)*	-.02 (.01)	-.05 (.02)**	-.02 (.01)
Agreeable	.12 (.02)**	.12 (.02)**	.13 (.02)**	.13 (.02)**	-.03 (.02)	-.05 (.02)**	-.07 (.02)**	-.04 (.02)*
Conscientious	.13 (.02)**	.13 (.02)**	.13 (.02)**	.13 (.02)**	-.07 (.02)**	-.09 (.02)**	-.10 (.02)**	-.08 (.02)**
Emotional Stability	.14 (.02)**	.15 (.02)**	.14 (.02)**	.14 (.02)**	-.22 (.02)**	-.22 (.02)**	-.23 (.02)**	-.22 (.020**
Openness	.03 (.02)	.03 (.02)	.02 (.02)	.03 (.02)	.01 (.02)	.02 (.02)	-.0002 (.02)	.02 (.02)
Adjusted R-squared	.2463	.2463	.2503	.2467	.1888	.1792	.1811	.1822
Observations	2,413	2,413	2,413	2,413	2,413	2,413	2,413	2,413

Notes: *$p < .05$, **$p < .01$ for 1-tailed test of a preregistered directional hypothesis. Otherwise, 2-tailed tests *p*-values reported**.** *Positive Mood* and *Negative Mood* were constructed as an average of the mood reports across the positive and negative mood dimensions assessed, respectively (as was preregistered).

TABLE 5: Hypothesis 4 tests—Psychopathy and Machiavellianism predictions on mood immediately *after* Trolley dilemma task

	Dependent Variable = *Positive Mood* *(Post-Trolley Dilemma task)*				**Dependent Variable = *Negative Mood* *(Post-Trolley Dilemma task)***			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Independent Variable	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**
constant	1.39 (.22)**	1.30 (.22)**	1.41 (.18)**	1.39 (.21)**	3.07 (.23)**	3.46 (.23)**	3.29 (.19)**	3.11 (.22)**
Psychopathy	.09 (.05)	---	---	---	.17 (.05)**	---	---	---
Machiav.	---	.09 (.04)*	---	---	---	.04 (.04)	---	---
Narcissism	---	---	.15 (.05)**	---	---	---	.19 (.05)**	---
Sadism	---	---	---	.09 (.05)*	---	---	---	.17 (.05)**
Age	.01 (.002)**	.01 (.002)**	.01 (.002)**	.01 (.002)**	-.001 (.002)	-.002 (.002)	-.001 (.002)	-.001 (.002)
Female (=1)	-.38 (.06)**	-.39 (.05)**	-.38 (.05)**	-.38 .06)**	.14 (.06)*	.10 (.06)	.12 (.06)*	.13 (.06)*
Extraversion	.19 (.02)**	.19 (.02)**	.17 (.02)**	.19 (.02)**	-.01 (.02)	-.01 (.02)	-.04 (.02)*	-.02 (.02)
Agreeable	.11 (.02)**	.11 (.02)**	.10 (.02)**	.11 (.02)**	.01 (.02)	-.01 (.02)	-.003 (.02)	.01 (.02)
Conscientious	.13 (.02)**	.13 (.02)**	.12 (.02)**	.13 (.02)**	-.09 (.02)**	-.10 (.02)**	-.11 (.02)**	-.09 (.02)**
Emotional Stability	.13 (.02)**	.13 (.12)**	.12 (.02)**	.13 (.02)**	-.21 (.02)**	-.21 (.02)**	-.22 (.02)**	-.21 (.02)**
Openness	.01 (.02)	.01 (.02)	-.003 (.02)	.01 (.02)	.01 (.02)	.01 (.02)	-.003 (.02)	.01 (.02)
Adj R-squared	.1936	.1943	.1960	.1939	.1047	.1008	.1065	.1050
Observations	2,413	2,413	2,413	2,413	2,413	2,413	2,413	2,413

Notes: *$p < .05$, **$p < .01$ for 1-tailed test of a preregistered directional hypothesis. Otherwise, 2-tailed tests *p*-values reported**.** *Positive Mood* and *Negative Mood* were constructed as an average of the mood reports across the positive and negative mood dimensions assessed, respectively (as was preregistered).

REFERENCES

Azizli, N., Atkinson, B. E., Baughman, H. M., Chin, K., Vernon, P. A., Harris, E., & Veselka, L. (2016). Lies and crimes: Dark Triad, misconduct, and high-stakes deception. *Personality and Individual Differences*, *89*, 34-39.

Back, M. D., Küfner, A. C., Dufner, M., Gerlach, T. M., Rauthmann, J. F., & Denissen, J. J. (2013). Narcissistic admiration and rivalry: disentangling the bright and dark sides of narcissism. *Journal of Personality and Social Psychology*, *105*(6), 1013.

Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, *121*(1), 154-161.

Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of mice, men, and trolleys: Hypothetical judgment versus real-life behavior in trolley-style moral dilemmas. *Psychological Science*, *29*(7), 1084-1093.

Brown, W. M., Hazraty, S., & Palasinski, M. (2019). Examining the dark tetrad and its links to cyberbulling. *Cyberpsychology, Behavior, and Social Networking*, *22*(8), 552-557.

Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences*, *67*, 97-102.

Campbell, J., Schermer, J. A., Villani, V. C., Nguyen, B., Vickers, L., & Vernon, P. A. (2009). A behavioral genetic study of the Dark Triad of personality and moral development. *Twin Research and Human Genetics*, *12*(2), 132-136.

Čopková, R., & Christenková, Z. (2021). The Effect of Dark Triad Traits on Decision-Making Styles. *Psychological Thought*, *14*(1), 74.

Crysel, L. C., Crosier, B. S., & Webster, G. D. (2013). The Dark Triad and risk behavior. *Personality and Individual Differences*, *54*(1), 35-40.

Curtis, G. J., Clare, J., Vieira, E., Selby, E., & Jonason, P. K. (2022). Predicting contract cheating intentions: Dark personality traits, attitudes, norms, and anticipated guilt and shame. *Personality and Individual Differences*, *185*, 111277.

DeScioli, P., Bruening, R., & Kurzban, R. (2011). The omission effect in moral cognition: Toward a functional explanation. *Evolution and Human Behavior*, *32*(3), 204-215.

Dickinson, D. L., & Masclet, D. (2021). Unethical Decision Making and Sleep Restriction: Experimental Evidence. *IZA Discussion Paper No. 14537*.

Dickinson, D. L., & Masclet, D. (2019). Using ethical dilemmas to predict antisocial choices with real payoff consequences: An experimental study. *Journal of Economic Behavior & Organization*, *166*, 195-215.

Dubreuil, B., & Grégoire, J. F. (2013). Are moral norms distinct from social norms? A critical assessment of Jon Elster and Cristina Bicchieri. *Theory and Decision*, *75*, 137-152.

Egan, V., Hughes, N., & Palmer, E. J. (2015). Moral disengagement, the dark triad, and unethical consumer attitudes. *Personality and Individual Differences*, *76*, 123-128.

Elster, J. (1989). Social norms and economic theory. *Journal of Economic Perspectives*, *3*(4), 99-117.

Foot, P. (1967). The problem of abortion and the doctrine of the double effect. *Oxford Review*, *5*.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*(1), 26.

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504-528.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105-2108.

Harrison, A., Summers, J., & Mennecke, B. (2018). The effects of the dark triad on unethical behavior. *Journal of Business Ethics*, *153*(1), 53-77.

Hershfield, H. E., Cohen, T. R., & Thompson, L. (2012). Short horizons and tempting situations: Lack of continuity to our future selves leads to unethical decision making and behavior. *Organizational Behavior and Human Decision Processes*, *117*(2), 298-310.

Jonason, P. K., Strosser, G. L., Kroll, C. H., Duineveld, J. J., & Baruffi, S. A. (2015). Valuing myself over others: The Dark Triad traits and moral and social values. *Personality and Individual Differences*, *81*, 102-106.

Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3) a brief measure of dark personality traits. *Assessment*, *21*(1), 28-41.

Kapoor, H., & Kaufman, J. C. (2022). Unbound: The relationship among creativity, moral foundations, and dark personality. *The Journal of Creative Behavior*, *56*(2), 182-193.

Karandikar, S., Kapoor, H., Fernandes, S., & Jonason, P. K. (2019). Predicting moral decision-making with dark personalities and moral values. *Personality and Individual Differences*, *140*, 70-75.

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.

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research*, *39*(2), 329-358.

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*(7), 1571-1582.

Lee, K., Ashton, M. C., Wiltshire, J., Bourdage, J. S., Visser, B. A., & Gallucci, A. (2013). Sex, power, and money: Prediction from the Dark Triad and Honesty–Humility. *European Journal of Personality*, *27*(2), 169-184.

March, E., & Marrington, J. Z. (2021). Antisocial and prosocial online behaviour: Exploring the roles of the dark and light triads. *Current Psychology*, 1-4.

Masclet, D., & Dickinson, D.L. (2019). Incorporating conditional morality into economic decisions. *IZA Discussion Paper No. 12782*.

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.

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. (2011). Measuring social value orientation. *Judgment and Decision making*, *6*(8), 771-781.

Nathanson, C., Paulhus, D. L., & Williams, K. M. (2006a). Predictors of a behavioral measure of scholastic cheating: Personality and competence but not demographics. *Contemporary Educational Psychology*, *31*(1), 97-122.

Nathanson, C., Paulhus, D. L., & Williams, K. M. (2006b). Personality and misconduct correlates of body modification and other cultural deviance markers. *Journal of Research in Personality*, *40*(5), 779-802.

Olsen, K. J., & Stekelberg, J. (2016). CEO narcissism and corporate tax sheltering. *The journal of the American Taxation Association*, *38*(1), 1-22.

Oluf, G. A., & Furnham, A. (2015). The relationship between bright-and dark-side personality traits. *Personality and Individual Differences*, *87*, 206-211.

O'Reilly, C. A., & Hall, N. (2021). Grandiose narcissists and decision making: Impulsive, overconfident, and skeptical of experts–but seldom in doubt. *Personality and Individual Differences*, *168*, 110280.

Palan, S., & Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22-27.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153-163.

Plouffe, R. A., Saklofske, D. H., & Smith, M. M. (2017). The assessment of sadistic personality: Preliminary psychometric evidence for a new measure. *Personality and Individual Differences*, *104*, 166-171.

Pollock, N. C., McCabe, G. A., Southard, A. C., & Zeigler-Hill, V. (2016). Pathological personality traits and emotion regulation difficulties. *Personality and Individual Differences*, *95*, 168-177.

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making*, *29*(5), 453-469.

Rauthmann, J. F., & Kolar, G. P. (2012). How "dark" are the Dark Triad traits? Examining the perceived darkness of narcissism, Machiavellianism, and psychopathy. *Personality and Individual Differences*, *53*(7), 884-889.

Rijsenbilt, A., & Commandeur, H. (2013). Narcissus enters the courtroom: CEO narcissism and fraud. *Journal of Business Ethics*, *117*(2), 413-429.

Sevi, B., Urganci, B., & Sakman, E. (2020). Who cheats? An examination of light and dark personality traits as predictors of infidelity. *Personality and Individual Differences*, *164*, 110126

Stöber, J. (2001). The Social Desirability Scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment*, *17*(3), 222.

Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality*, *60*(2), 225-251.

Van Scotter, J. R., & Roglio, K. D. D. (2020). CEO bright and dark personality: Effects on ethical misconduct. *Journal of Business Ethics*, *164*(3), 451-475.

Walker, S. A., Olderbak, S., Gorodezki, J., Zhang, M., Ho, C., & MacCann, C. (2022). Primary and secondary psychopathy relate to lower cognitive reappraisal: A meta-analysis of the Dark Triad and emotion regulation processes. *Personality and Individual Differences*, *187*, 111394.

Zhao, H., Zhang, H., & Xu, Y. (2016). Does the dark triad of personality predict corrupt intention? The mediating role of belief in good luck. *Frontiers in Psychology*, *7*, 608.

Zeigler-Hill, V., & Vonk, J. (2015). Dark personality features and emotion dysregulation. *Journal of Social and Clinical Psychology*, *34*(8), 692.

Zerbe, W. J., & Paulhus, D. L. (1987). Socially desirable responding in organizational behavior: A reconception. *Academy of Management Review*, *12*(2), 250-264.

Zuo, S., Wang, F., Xu, Y., Wang, F., & Zhao, X. (2016). The fragile but bright facet in the Dark Gem: Narcissism positively predicts personal morality when individual's self-esteem is at low level. *Personality and Individual Differences*, *97*, 272-276.

Dark versus light personality types and moral choice
David L Dickinson

ONLINE SUPPLEMENTAL APPENDIX A
(Estimation Results not shown in main text)

TABLE A1: Results from estimations producing Figure 3 (main text)

Dependent Var = *SVO Angle*	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Independent Variable	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**
constant	23.01 (2.30)**	32.22 (2.29)**	17.88 (1.85)**	23.35 (2.19)**	2.85 (2.06)	5.43 (2.02)**	9.39 (1.79)**
Psychopathy	-2.76 (.49)**	---	---	---	---	---	---
Machiavellianism	---	-4.42 (.39)**	---	---	---	---	---
Narcissism	---	---	-2.26 (.47)**	---	---	---	---
Sadism	---	---	---	-3.11 (.48)**	---	---	---
Kantianism	---	---	---	---	3.76 (.43)**	---	---
Humanism	---	---	---	---	---	3.27 9.46)**	---
Faith in Humanity	---	---	---	---	---	---	2.24 (.36)**
Age	.02 (.02)	.01 (.02)	.02 (.02)	.01 (.02)	.02 (.02)	.05 (.02)**	.04 (.02)*
Female (=1)	.92 (.58)	1.03 (.56)	1.40 (.57)*	1.01 (.57)	1.54 (.56)**	1.27 (.57)*	1.77 (.57)**
Extraversion	.25 (.17)	.14 (.16)	.54 (.18)**	.24 (.16)	.30 (.16)	-.02 (.17)	-.06 (.17)
Agreeable	1.15 (.24)**	.80 (.23)**	1.48 (.23)**	1.07 (.24)**	1.07 (.23)**	1.02 (.24)**	1.05 (.24)**
Conscientious	-.44 (.22)*	-.31 (.21)	-.16 (.21)	-.39 (.21)	-.43 (.21)*	-.32 (.21)	-.18 (.21)
Emotional Stability	-.67 (.19)**	-.60 (.19)**	-.56 (.20)**	-.63 (.19)**	-.60 (.19)**	-.54 (.19)**	-.74 (.19)**
Openness	.22 (.22)	.08 (.21)	.31 (.22)	-.09 (.22)	-.001 (.33)	-.07 (.22)	.18 (.22)
Adj. R-squared	.0427	0790	0391	.0469	.0601	.0503	.0458
Observations	2,413	2,413	2,413	2,413	2,413	2,413	2,413

Notes: *$p < .05$, **$p < .01$ for the 1-tailed test of the preregistered hypotheses (all other p-values reported for 2-tailed tests). The coefficient plots in Fig. 3 (main text) show the point estimates and confidence intervals on the specific trait's impact on *SVO Angle* (these estimates are shaded)

TABLE A2: Results from estimations producing Figure 4 (main text)

Dependent Variable = *HEADS*	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Independent Variable	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**
constant	5.84 (.30)**	5.66 (.30)**	6.04 (.24)**	5.85 (.29)**	6.39 (.27)**	6.37 (.26)**	6.28 (.23)**
Psychopathy	.11 (.06)*	---	---	---	---	---	---
Machiavellianism	---	.13 (.05)**	---	---	---	---	---
Narcissism	---	---	.08 (.06)	---	---	---	---
Sadism	---	---	---	.11* (.06)	---	---	---
Kantianism	---	---	---	---	-.07 (.06)	---	---
Humanism	---	---	---	---	---	-.07 (.06)	---
Faith in Humanity	---	---	---	---	---	---	-.04 (.05)
Age	-.01 (.002)**	-.01 (.003)**	-.01 (.002)**	-.01 (.002)**	-.01 (.003)**	-.01 (.002)**	-.01 (.002)**
Female (=1)	-.11 (.08)	-.12 (.07)	-.13 (.07)	-.12 (.07)	-.13 (.07)	-.13 (.07)	-.14 (.07)
Extraversion	.008 (.02)	.01 (.02)	-.002 (.02)	.01 (.02)	.01 (.02)	.02 (.02)	.02 (.02)
Agreeable	.002 (.03)	.008 (.03)	-.01 (.03)	.004 (.03)	-.01 (.03)	-.002 (.03)	-.004 (.03)
Conscientious	-.01 (.03)	-.02 (.03)	-.02 (.03)	-.01 (.03)	-.02 (.03)	-.02 (.03)	-.02 (.03)
Emotional Stability	-.004 (.03)	-.01 (.03)	-.01 (.03)	-.01 (.03)	-.01 (.03)	-.01 (.03)	-.003 (.03)
Openness	.01 (.03)	.02 (.03)	.008 (.03)	.02 (.03)	.02 (.03)	.02 (.03)	.01 (.03)
Adj. R-squared	.0127	.0140	.0123	.0129	.0122	.0121	.0119
Observations	2,413	2,413	2,413	2,413	2,413	2,413	2,413

Notes: *$p < .05$, **$p < .01$ for the 1-tailed test of the preregistered hypotheses (all other p-values reported for 2-tailed tests). The coefficient plots in Fig. 3 (main text) show the point estimates and confidence intervals on the specific trait's impact on *HEADS* reported in the Coin Flip task (these estimates are shaded)

TABLE A3: Utilitarian choice as predicted by specific personality traits (see also Table 3 in the main text)

Dependent Variable = *Proportion Utilitarian Choices*	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Independent Variable	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**
constant	.93 (.04)**	.90 (.04)**	.90 (.03)**	.91 (.03)**	.88 (.03)**	.83 (.03)**	.85 (.03)**
Psychopathy	-.01 (.01)	---	---	---	---	---	---
Machiavellianism	---	-.001 (.006)	---	---	---	---	---
Narcissism	---	---	-.004 (.008)	---	---	---	---
Sadism	---	---	---	-.004 (.008)	---	---	---
Kantianism	---	---	---	---	.006 (.007)	---	---
Humanism	---	---	---	---	---	.03 (.007)**	---
Faith in Humanity	---	---	---	---	---	---	.02 (.006)**
Age	-.0001 (.0003)	-.0000 (.0003)	-.0001 (.0003)	-.0001 (.0003)	-.0000 (.0003)	-.0001 (.0003)	-.0001 (.0003)
Female (=1)	-.01 (.01)	-.01 (.01)	-.01 (.01)	-.01 (.01)	-.01 (.01)	-.01 (.01)	-.01 (.01)
Extraversion	.003 (.003)	.003 (.003)	.003 (.003)	.003 (.003)	.003 (.003)	.001 (.003)	.001 (.003)
Agreeable	.007 (.004)	.009 (.004)*	.009 (.004)*	.008 (.004)*	.008 (.004)*	.004 (.003)	.004 (.004)
Conscientious	-.001 (.003)	-.0003 (.003)	-.0002 (.003)	-.0005 (.003)	-.001 (.003)	-.001 (.003)	.0000 (.003)
Emotional Stability	-.006 (.003)*	-.006 (.003)*	-.006 (.003)	-.006 (.003)*	-.006 (.003)*	-.005 (.003)	-.007 (.003)*
Openness	-.001 (.003)	-.001 (.003)	-.0001 (.0003)	-.0001 (.0003)	-.001 (.003)	-.003 (.003)	-.0004 (.003)
Adj. R-squared	.0010	.0004	.0004	0005	.0006	.0057	.0061
Observations	2,413	2,413	2,413	2,413	2,413	2,413	2,413

Notes: *$p < .05$, **$p < .01$ for 1-tailed test of a preregistered directional hypothesis. Some seemingly different significance level indicators are due to rounding of coefficient and/or standard errors reported here. Otherwise, 2-tailed tests *p*-values reported. The *Proportion Utilitarian Choices* variables takes on the value of 0, .25, .5, .75, or 1 and represents the proportion of utilitarian choices made across the 4 Trolley scenarios with a unique utilitarian choice (i.e., (1:1) and (5:5) Trolley scenario choices excluded for this analysis).

TABLE A4: Results from estimations producing Figure 5 (main text)

Dependent Variable = *NET-Positive mood change (post-Trolley)*	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Independent Variable	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**
constant	.18 (.30)	.13 (.30)	.91 (.24)**	.65 (.29)*	1.14 (.27)**	1.19 (.26)**	1.05 (.23)**
Psychopathy	.23 (.06)**	---	---	---	---	---	---
Machiavellianism	---	.19 (.05)**	---	---	---	---	---
Narcissism	---	---	.01 (.06)	---	---	---	---
Sadism	---	---	---	.09 (.06)	---	---	---
Kantianism	---	---	---	---	-.07 (.05)	---	---
Humanism	---	---	---	---	---	-.10 (.06)	---
Faith in Humanity	---	---	---	---	---	---	-.06 (.04)
Age	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**
Female (=1)	-.33 (.08)**	-.37 (.07)**	-.39 (.07)**	-.38 (.08)**	-.39 (.07)**	-.38 (.07)**	-.40 (.07)**
Extraversion	.000 (.02)	.01 (.02)	.01 (.02)	.01 (.02)	.01 (.02)	.01 (.02)	.01 (.02)
Agreeable	-.06 (.03)	-.06 (.03)	-.10 (.03)**	-.08 (.03)*	-.09 (.03)**	-.08 (.03)*	-.08 (.03)*
Conscientious	.03 (.03)	.01 (.03)	.01 (.03)	.01 (.03)	.01 (.03)	.01 (.03)	.01 (.03)
Emotional Stability	-.03 (.03)	-.03 (.03)	-.03 (.03)	-.03 (.03)	-.03 (.03)	-.03 (.03)	-.03 (.03)
Openness	-.03 (.03)	-.02 (.03)	-.02 (.03)	-.02 (.03)	-.02 (.03)	-.02 (.03)	-.02 (.03)
Adj. R-squared	.0314	.0318	.0263	.0271	.0269	.0274	.0269
Observations	2,413	2,413	2,413	2,413	2,413	2,413	2,413

Note: *$p < .05$, **$p < .01$ for 2-tailed test. The dependent variable, *Net Positive Mood Change*, is defined as the difference between the positive and negative mood change from baseline to post-Trolley task. Thus, an increase in the value of this variable indicates that one's mood turned more positive (or less negative) after completing the Trolley dilemma task.

TABLE A5: Personality trait impacts on Trolley dilemma immorality-Acts of Omission

Dependent Variable = *Immoral Act of Omission* (=1)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Independent Variable	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**
constant	-.05 (03)	.004 (.03)	.001 (.02)	-.03 (.03)	.03 (.03)	.06 (.03)*	.04 (.02)
Psychopathy	.02 (.01)**	---	---	---	---	---	---
Machiavellianism	---	.001 (.005)	---	---	---	---	---
Narcissism	---	---	.004 (.006)	---	---	---	---
Sadism	---	---	---	.013 (.006)*	---	---	---
Kantianism	---	---	---	---	-.009 (.005)	---	---
Humanism	---	---	---	---	---	-.02 (.006)**	---
Faith in Humanity	---	---	---	---	---	---	-.01 (.005)**
Age	.0001 (.0002)	.0000 (.0002)	.0000 (.0002)	.0001 (.0002)	.0000 (.0002)	-.0000 (.0002)	-.0000 (.0002)
Female (=1)	.001 (.01)	-.003 (.01)	-.003 (.01)	-.001 (.007)	-.003 (.007)	-.001 (.007)	-.004 (.007)
Extraversion	.001 (.002)	.001 (.003)	.001 (.002)	.001 (.002)	.001 (.002)	.002 (.002)	.003 (.002)
Agreeable	-.03 (.003)	-.01 (.003)	-.01 (.003)	-.003 (.003)	-.004 (.003)	-.002 (.003)	-.002 (.003)
Conscientious	.004 (.003)	.002 (.003)	.002 (.003)	.003 (.003)	.003 (.003)	.003 (.003)	.002 (.003)
Emotional Stability	.004 (.002)	.004 (.002)	.004 (.002)	.004 (.002)	.004 (.002)	.004 (.002)	.005 (.002)*
Openness	.002 (.003)	.002 (.003)	.002 (.003)	.002 (.003)	.003 (.003)	.005 (.003)	.002 (.003)
Adj. R-squared	.0042	.0014	.0015	.0034	.0025	.0060	.0058
Observations	2,413	2,413	2,413	2,413	2,413	2,413	2,413

Notes: *$p < .05$, **$p < .01$ for 1-tailed test of a preregistered directional hypothesis on sadism (all others p-values indicated are for the 2-tailed test). Coefficient estimates on *sadism* are those reported in the main text, Figure 6. Some seemingly differing significance level indications are due to rounding used for brevity in the table.

TABLE A6: Personality trait impacts on Trolley dilemma immorality—Acts of Commission

Dependent Variable = *Immoral Act of Commission* (=1)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Independent Variable	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**	**Coef (st. error)**
constant	-.14 (.05)**	-.02 (.05)	.007 (.04)	-.04 (.05)	.11 (.05)*	.07 (.05)	.05 (.04)
Psychopathy	.06 (.01)**	---	---	---	---	---	---
Machiavellianism	---	.02 (.01)*	---	---	---	---	---
Narcissism	---	---	.03 (.01)**	---	---	---	---
Sadism	---	---	---	.04 (.01)**	---	---	---
Kantianism	---	---	---	---	-.02 (.01)	---	---
Humanism	---	---	---	---	---	-.02 (.01)	---
Faith in Humanity	---	---	---	---	---	---	.01 (.01)
Age	.001 (.0004)**	.001 (.0004)*	.001 (.0004)*	.001 (.0004)*	.001 (.0004)*	.001 (.0004)*	.001 (.0004)*
Female (=1)	-.02 (.01)	-.03 (.01)*	-.03 (.01)*	-.03 (.01)*	-.03 (.01)*	-.03 (.01)**	-.03 (.01)**
Extraversion	.01 (.003)**	.02 (.004)**	.01 (.004)**	.02 (.004)**	.02 (.004)**	.02 (.004)**	.02 (.004)**
Agreeable	.02 (.006)**	.01 (.006)	.01 (.004)	.01 (.006)*	.01 (.01)	.006 (.006)	.004 (.006)
Conscientious	.005 (.005)	.0004 (.005)	-.001 (.005)	.002 (.005)	.001 (.005)	.0001 (.005)	.0001 (.005)
Emotional Stability	-.005 (.005)	-.006 (.004)	-.007 (.005)	-.006 (.005)	-.006 (.005)	-.006 (.005)	-.006 (.005)
Openness	-.01 (.005)	-.008 (.005)	-.01 (.005)	-.01 (.005)	-.008 (.005)	-.008 (.005)	-.008 (.005)
Adj. R-squared	.0205	.0108	.0120	.0129	.0099	.0088	.0090
Observations	2,413	2,413	2,413	2,413	2,413	2,413	2,413

Notes: *$p < .05$, **$p < .01$ for 1-tailed test of a preregistered directional hypothesis on sadism (all others p-values indicated are for the 2-tailed test). Coefficient estimates on *sadism* are those reported in the main text, Figure 6. Some seemingly differing significance level indications are due to rounding used for brevity in the table.

TABLE A7: H6 test (see main text, Figure 7)

	Dependent Variable = *HEADS* reported							
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Independent Variable	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**	**Coef (st. err)**
constant	5.55 (.03)**	6.13 (.11)**	6.19 (.21)**	6.13 (.22)**	5.56 (.03)**	6.14 (.11)**	6.20 (.21)**	6.13 (.22)**
Act of Omission	-.13 (.21)	-.12 (.21)	-.12 (.21)	-.16 (.21)	---	---	---	---
Act of Commission	---	---	---	---	-.09 (.11)	-.13 (.11)	-.13 (.11)	-.15 (.11)
Age	---	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**	---	-.01 (.002)**	-.01 (.002)**	-.01 (.002)**
Female (=1)	---	-.14 (.07)**	-.14 (.07)	-.11 (.07)	---	-.14 (.07)*	-.14 (.07)	-.12 (.078)
Extraversion	---	---	.01 (.02)	.01 (.02)	---	---	.01 (.02)	.01 (.02)
Agreeable	---	---	-.02 (.03)	.02 (.03)	---	---	-.01 (.03)	.02 (.03)
Conscientious	---	---	-.02 (.03)	-.01 (.03)	---	---	-.02 (.03)	-.01 (.03)
Emotional Stability	---	---	-.004 (.03)	-.01 (.03)	---	---	-.01 (.03)	-.01 (.03)
Openness	---	---	.01 (.03)	.02 (.03)	---	---	.01 (.03)	.01 (.03)
NetLight	---	---	---	-.12 (.05)*	---	---	---	-.12 (.05)*
Adjusted R-squared	-.0002	.0130	.0117	.0137	-.0002	.0134	.0121	.0142
Observations	2,463	2,413	2,413	2,413	2,463	2,413	2,413	2,413

Notes: *$p < .05$, **$p < .01$ for 2-tailed test. Coefficient estimates of the variables identifying the hypothesis test are shaded above.

TABLE A8: Instrumental Variables **(**Exploratory) Analysis of link between mood response and moral choice

	1st Stage regression		
Independent variable	Dependent Variable ***NET-Positive mood change (post-Trolley)***		
Constant	.024 (.072)		
NetLight	-.240 (.039)**		
Observations	2463		
R-squared	.0148		
	2nd Stage regression Instruments = *NetLight* Instrumented Variable = *NET-Positive mood change (post-Trolley)*		
Independent variable	Dependent Variable ***Immoral Omission***	Dependent Variable ***Immoral Commission***	Dependent Variable ***HEADS***
Constant	.045 (.008)**	.118 (.013)**	5.777 (.079)**
NET-Positive mood change (post-Trolley)	.056 (.018)**	.075 (.032)*	.585 (.191)**
Observations	2464	2463	2463
Wald X^2 (p-value)	9.47 (p = .0021)	5.37 (p = .0204)	9.43 (p = .0021)

Notes: *p < .05, **p < .01 for the 2-tailed test of the coefficient estimates. This table highlights that the increased positive mood post-Trolley explained by one's darkness of personality measure predicts immoral choice (hypothetical and consequential) in the 2nd stage estimation equations.