A Dose of Ruthlessness: Interpersonal Moral Judgment Is Hardened by the Anti-Anxiety Drug Lorazepam

Adam M. Perkins, Ania M. Leonard, Kristin Weaver, Jeffrey A. Dalton, Mitul A. Mehta, Veena Kumari, and Steven C. R. Williams
King’s College London

Ulrich Ettinger
University of Bonn

Neuroimaging data suggest that emotional brain systems are more strongly engaged by moral dilemmas in which innocent people are directly harmed than by dilemmas in which harm is remotely inflicted. In order to test the possibility that this emotional engagement involves anxiety, we investigated the effects of 1 mg and 2 mg of the anti-anxiety drug lorazepam on the response choices of 40 healthy volunteers (20 male) in moral-personal, moral-impersonal, and nonmoral dilemmas. We found that lorazepam caused a dose-dependent increase in participants’ willingness to endorse responses that directly harm other humans in moral-personal dilemmas but did not significantly affect response choices in moral-impersonal dilemmas or nonmoral dilemmas. Within the set of moral-personal dilemmas that we administered, lorazepam increased the willingness to harm others in dilemmas where harm was inflicted for selfish reasons (dubbed low-conflict dilemmas) as well as responses to dilemmas where others were harmed for utilitarian reasons (i.e., for the greater good, dubbed high-conflict dilemmas). This suggests that anxiety exerts a general inhibitory effect on harmful acts toward other humans regardless of whether the motivation for those harmful acts is selfish or utilitarian. Lorazepam is also a sedative drug, but we found that lorazepam slowed decision times equally in all 3 dilemma types. This finding implies that its specific capacity to increase ruthlessness in moral-personal dilemmas was not a confound caused by sedation.

Keywords: moral judgment, anxiety, utilitarianism, ruthlessness, psychopathy

Supplemental materials: http://dx.doi.org/10.1037/a0030256.supp

Philosophers have historically emphasized the governing role of rational thought in human moral judgments (Robinson, 1971), whereas more recent psychological analyses suggest that emotion is also important in adaptive moral functioning (e.g., Damasio, 1994). Neuroimaging data suggest that both processes are employed in moral reasoning but the degree of involvement depends upon the type of moral dilemma, as emotion-related brain areas are more strongly engaged by dilemmas in which innocent people are directly harmed (e.g., “Would you kill one healthy person so his organs can be used to save five terminally ill people?”) than by dilemmas in which harm is remotely inflicted (e.g., “Would you operate a switch that saves five railway workers from being killed by a runaway train at the expense of running over and killing another worker?”; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). It has been suggested that the prospect of directly harming others, even for the greater good, generates an aversive emotional “gut” reaction that must then be overcome by controlled cognitive processes in order for a directly harmful action to be performed (Greene, Nystrom, Engell, Darley, & Cohen, 2004). This notion has been supported by the finding that cognitive load selectively increased reaction times for responses that required direct harm to be done to other humans for the greater good (Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008). Further support has come from the finding that this type of response increased during experimental manipulations that increased cognitive reflection in participants (Paxton, Ungar, & Greene, 2012) but decreased when participants were required to respond quickly (Suter & Hertwig, 2011).
However, the human emotional repertoire encompasses various types of positive and negative affect (e.g., Ekman, 1999; Perkins, Inchley-Mort, Pickering, Corr, & Burgess, 2012), making it useful to narrow down specifically which part of the emotional spectrum is responsible for inhibiting behavior that directly harms other people. Indirect evidence points to an inhibitory role for negative emotion. For example, Valdesolo and DeSteno (2006) found that experimentally increasing mirth-related positive affect in participants increased their tendency to endorse harmful acts in moral-personal dilemmas, implying that negative emotion should conversely play an inhibitory role in such acts. Similarly, a single 30-mg dose of citalopram (a selective serotonin reuptake inhibitor drug that is used to treat a wide range of affective illnesses, including anxiety disorders; Wade, Lepola, Koponen, Pedersen, & Pedersen, 1997) decreased the likelihood that participants would endorse harmful acts in moral-personal dilemmas (Crockett, Clark, Hauser, & Robbins, 2010). This finding seems contradictory, but it should be noted that single doses of drugs that block the reuptake of serotonin increase the intensity of threat avoidance behavior in rodents (Griebel, Blanchard, Agnes, & Blanchard, 1995), implying that they have a short-term boosting effect on negative emotional responses. The inhibiting effect of citalopram on harmful acts that was found by Crockett et al. (2010) may therefore legitimately be interpreted as a product of this temporary boosting effect on negative emotion caused by citalopram.

A Role for Anxiety in Moral Judgment?

In order to refine further our understanding of the emotions underlying moral judgment, research is required that can more directly test which negative emotion interferes with endorsing directly harmful acts in moral-personal dilemmas. Anxiety is a plausible candidate for this emotion, as clinical observations show that interpersonal situations are particularly strong eliciting stimuli for anxiety, in humans at least (American Psychiatric Association, 1994). Additionally, it has been found with the moral dilemma task that low-anxiety psychopaths are more likely than either control participants or high-anxiety psychopaths to endorse directly harmful behavior in moral-personal dilemmas. In contrast, the two groups of psychopaths were both more likely than control participants to endorse harmful impersonal acts that cause indirect or remote harm to others (Koenigs, Kruepke, Zeier, & Newman, 2012).

Both clinical observations and criterion group studies of psychopaths therefore suggest that anxiety plays a key role in modulating preferences for directly harmful acts in moral-personal dilemmas, but these are correlational data that cannot test for a causal role for anxiety in moral decision making. One research strategy that has been regarded for several decades as being able to test causation with regard to anxiety is to measure the behavioral effects of anti-anxiety drugs:

It is well established, for example, that the benzodiazepines (and other classes of anxiolytic drug) increase approach to locations or stimuli that have been associated with painful electric shock, as well as willingness to perform an operant, e.g., barpressing, that is sometimes followed by electric shock (Gray, 1977). Indeed, the lack of effects of this kind in animal experiments would be sufficient grounds to cast serious doubt upon the claim that a given drug is an anxiolytic. (Sartory, MacDonald, & Gray, 1990, p. 273)

If it is true that anxiety inhibits the endorsement of directly harmful acts in moral-personal dilemmas, then administering an anti-anxiety drug should systematically increase the willingness of participants to endorse directly harmful acts in moral-personal dilemmas without significantly affecting response choices in other types of dilemmas. To address this gap in the literature, we investigated the effects, versus placebo, of 1 mg and 2 mg lorazepam, a high-potency 3-hydroxy benzodiazepine commonly used for the management of anxiety disorders and short-term relief from anxiety in other conditions (Gould, Otto, Pollack, & Yap, 1997), on the response choices of healthy participants in moral-personal, moral-impersonal, and nonmoral dilemmas.

Specificity of Anti-Anxiety Drug Effects

In a pharmacological experiment of this type, one important issue is whether anti-anxiety drugs cause their effects by altering anxiety or through general emotional and/or cognitive alteration. For example, such drugs might achieve their clinically useful anti-anxiety effects via a scattergun effect that dampens emotional responses of all kinds. If this were true, it would not necessarily diminish their therapeutic value, but when these drugs are used as experimental probes, as in the present study, such a general effect would make it difficult to say what type of emotion is important in inhibiting directly harmful acts during moral reasoning.

A range of objective evidence concerning the behavioral effects of anti-anxiety drugs suggests that they do not have a scattergun effect on emotion; instead, their effects are limited to dampening emotion produced by threat. For example, a substantial rodent literature shows that anti-anxiety drugs reduce perceptions of threat intensity, temporarily shifting the individual down a gradient of defensiveness that equates in physical terms to moving the animal further from the threat source and consequently diminishing threat-related emotional activation (i.e., anxiety; Blanchard, Griebel, & Blanchard, 2003). Research into the effects on threat perception of anti-anxiety drugs in humans is less developed than into those in rodents, but early signs are that they are similar. For example, benzodiazepine drugs reduce threat-potentiated startle but have no effect on baseline startle (e.g., Patrick, Berthot, & Moore, 1996; RibA et al., 2001) or on pleasure-attenuated startle (see Thornton et al.’s abstract in “Abstracts 25th Annual Meeting, British Psychophysiology Society,” 1999, p. 141). Congruent with these human startle data, the anti-anxiety drug diazepam has been found to reduce self-reported fear in snake phobics who were required to approach a live snake (Sartory et al., 1990).

In a similar vein, our research group has recently found that 1 mg of lorazepam significantly reduced the intensity of avoidance behavior in response to threat stimuli capable of inflicting a 115 dB white noise burst in healthy human participants. This finding suggests that this drug made the threat stimuli seem less dangerous and hence less deserving of intense avoidance reactions (Perkins et al., 2009). These data together suggest that anti-anxiety drugs in humans, as in rodents, reduce perceptions of threat intensity and thus dampen negative emotion elicited by aversive situations. Precisely how this threat-related negative emotion should be labeled is a matter of debate, considering the long-standing confusion over fear–anxiety separability (e.g., Geer, 1965). Because lorazepam is usually labeled as an anti-anxiety drug (Gould et al., 1997), in the present study we chose to label it as anxiety, while accepting that the threat-related emotion might also be labeled as...
worry, apprehension, concern, fear, or a range of other negative emotion descriptors that plausibly describe one’s affective reaction to aversive situations. Our primary hypothesis in this study was therefore that lorazepam would significantly increase in a dose-dependent manner our participants’ willingness to endorse harmful (i.e., ruthless) response choices in moral-personal dilemmas, in accordance with the notion that it is an anti-anxiety drug. Anxiety (as we label it here) hypothetically plays a role in inhibiting harmful behavior toward others in dilemmas of that type, behavior that we would ordinarily find aversive.

In addition to its anti-anxiety effects, lorazepam is known to cause sedation (e.g., Curran, Poovooonsuk, Dalton, & Lader, 1998; Mintzer & Griffiths, 2003; Vermeeren et al., 1995). This makes it necessary to verify that any changes in moral judgment caused by lorazepam are not merely the result of sedation. To address this concern we measured the effects of lorazepam on dilemma decision time, as an objective proxy measure of sedation. Our rationale for this step was that lorazepam and other sedative benzodiazepines are known to impair psychomotor performance and increase reaction time (Smiley, 1987; Van Ruitenbeek, Vermeeren, & Riedel, 2010). We therefore hypothesized that lorazepam would significantly increase decision time in all three dilemma types and that the change in decision time caused by lorazepam would not be significantly associated with changes that lorazepam produces in response choices to moral-personal dilemmas.

**Selfish Versus Utilitarian Ruthlessness**

Recent research (Greene et al., 2008; Koenigs et al., 2007) has shown that the set of moral-personal dilemmas used by Greene et al. (2001) can be divided into two conceptually distinct groups that have been dubbed low-conflict and high-conflict (see their supplemental materials). These two subsets of moral-personal dilemmas both ask participants whether they would endorse actions that cause direct harm to innocent people, but they differ in the extent to which those harmful acts can be justified as being for the greater good. This is an important philosophical distinction because integration of research using the dilemma task with the moral philosophy literature hinges on the difference between harming others for the greater good and harming others for less noble reasons. This former stance is typically known as utilitarianism (Mill, 1861/1998) and contrasts with the deontological view of moral reasoning, which condemns acts when they violate certain rights or duties, even if such acts ultimately promote the greater good (Kant, 1785/1959).

In low-conflict moral-personal dilemmas, the harmful acts are primarily selfish in their motivation, with little or no discernible benefit for the greater good (e.g., You are a young architect visiting one of your construction sites with your boss. Your boss is a despicable individual who makes everyone around him miserable including you. It occurs to you that if you were to push him off of the building you are inspecting he would fall to his death and everyone would think it was an accident. Is it appropriate for you to push your boss off of the building in order to get him out of your life? See Greene et al., 2001). In contrast, in moral-personal dilemmas categorized as high conflict, the harmful act clearly benefits the greater good and is thus truly utilitarian (e.g., You are the captain of a military submarine travelling underneath a large iceberg. An on board explosion has caused you to lose most of your oxygen supply and has injured one of your crew who is quickly losing blood. The injured crew member is going to die from his wounds no matter what happens. The remaining oxygen is not sufficient for the entire crew to make it to the surface. The only way to save the other crew members is to shoot dead the injured crew member so that there will be just enough oxygen for the rest of the crew to survive. Is it appropriate for you to kill the fatally injured crew member in order to save the lives of the remaining crew members? See Greene et al., 2001).

As an exploratory extension to our study we therefore sought to determine whether these two sets of moral-personal dilemmas would be differentially affected by lorazepam, with a view to determining if utilitarianism is underpinned by a lack of anxiety or whether associations between utilitarianism and low anxiety are a function of a general inhibiting effect of anxiety on actions that directly harm others, irrespective of whether such actions serve the greater good. If the former hypothesis is true, lorazepam should increase harmful (i.e., ruthless) responding in the high-conflict (i.e., utilitarian) dilemmas only, whereas if the latter is true, lorazepam should increase harmful responding in both the low-conflict and the high-conflict moral-personal dilemmas.

**Method**

**Participants**

Forty healthy volunteers (20 male) between 19 and 35 years of age (Mean = 24.80, SD = 4.08) were recruited via an e-mail advertisement sent to the staff and students of King’s College London. Participants were assessed by a psychiatrist before admission to the study to ensure they were in good physical and mental health. Their well-being was additionally verified before and after each experimental session. The study was approved by the King’s College Hospital Research Ethics Committee, and all participants gave written informed consent.

**Stimuli**

Participants were presented with a total of 54 different dilemmas during the course of the study, with 18 dilemmas being presented in each of the three testing sessions, by means of bespoke software programmed in VB.net. The dilemmas used in this study are listed in the supplementary information and were used in an earlier neuroimaging study of moral judgment in which they were sorted into moral-personal, moral-impersonal, and nonmoral categories by independent judges (Greene et al., 2001), with six dilemmas of each type being completed in each session. The phrasing of the moral dilemmas forced participants to make a clear choice between harming others or not. The phrasing of the nonmoral dilemmas forced participants to make a clear choice between rational and nonrational responses. Examples are as follows (Greene et al., 2001):

**Moral-personal.** You are a doctor. You have five patients, each of whom is about to die due to a failing organ of some kind. You have another patient who is healthy. The only way that you can save the lives of the first five patients is to transplant five of this young man’s organs (against his will) into the bodies of the other five patients. If you do this, the young man will die, but the
other five patients will live. Is it appropriate for you to perform this transplant in order to save five of your patients?  

**Moral-impersonal.** You are at the wheel of a runaway trolley quickly approaching a fork in the tracks. On the tracks extending to the left is a group of five railway workmen. On the tracks extending to the right is a single railway workman. If you do nothing the trolley will proceed to the left, causing the deaths of the five workmen. The only way to avoid the deaths of these workmen is to hit a switch on your dashboard that will cause the trolley to proceed to the right, causing the death of the single workman. Is it appropriate for you to hit the switch in order to avoid the deaths of the five workmen?  

**Nonmoral.** You are a farm worker driving a turnip-harvesting machine. You are approaching two diverging paths. By choosing the path on the left you will harvest 10 bushels of turnips. By choosing the path on the right you will harvest 20 bushels of turnips. If you do nothing your turnip-harvesting machine will turn to the left. Is it appropriate for you to turn your turnip-picking machine to the right in order to harvest 20 bushels of turnips instead of 10?  

**Procedure**

Each participant was tested three times, with the sessions being at least a week apart to allow for adequate drug washout, under placebo (50 mg ascorbic acid), 1 mg lorazepam, and 2 mg lorazepam in a randomized order. Drugs were administered with 300 ml of still water. After a 2-hr wait for drug metabolism to occur, participants completed six dilemmas from each of the three categories in each testing session. Dilemmas were arranged in a pseudo-random order to enhance unpredictability. No dilemma was presented twice to the same participant during the course of the experiment. Dilemmas were administered with computer software that recorded each response choice as well as the time taken to make a response, within a temporal resolution of approximately 10 milliseconds. The participants were seated alone in front of a computer monitor. Each dilemma was displayed singly on three consecutive slides, two displaying the dilemma text (split into two parts) and one requiring the participant to make a yes/no choice using the arrow keys on the computer keyboard (left arrow key = yes; right arrow key = no). The maximum time to read each slide describing the dilemma was 200 s and the maximum response time was 160 s, making the task essentially self-paced. The participants controlled the transition from slide to slide using the right arrow key, but once they moved on to the next screen they could not go back to the previous screen. To ensure that the right arrow key was not pressed accidentally when making a response, a 2-s safety delay, during which no further selections could be made, was active after the slide change.  

**Coding of Dilemma Response Choices**

To allow within-subjects comparison of drug effects on response choices between dilemma types, we scored moral dilemma response choices of either type using the same scheme; namely, that each endorsement of a harmful action added a point to the participant’s ruthlessness score for that dilemma type. Thus, higher scores indicate a greater willingness to endorse harmful choices (see supplemental materials for the complete list of dilemmas and scoring scheme used in this experiment). The control dilemmas contained no moral content and entailed no harm to anyone: They were therefore scored on a rational basis, with each endorsement of a rational action adding a point to their score. Scores for rationality in the nonmoral dilemmas were regarded for the purposes of statistical analysis as conceptually equivalent to ruthlessness scores for the moral dilemmas, as both rational and ruthless responses in the present context hypothetically denote nonmotional reactions.  

**Decision Times**

Dilemma decision times were scored automatically by the computer software that administered the task.  

**Statistical Analysis**

Statistical analyses were performed with the 18th edition of the Statistical Package for the Social Sciences (SPSS 18). Response choice data (frequency of ruthless/rational responses) were analyzed with repeated-measures analyses of variance (ANOVA) with the within-subjects factors dilemma type (moral-personal, moral-impersonal, nonmoral) and drug (placebo, 1 mg lorazepam, 2 mg lorazepam). Differences in lorazepam effects on response choices between the low-conflict and high-conflict moral-personal dilemmas were analyzed with repeated-measures ANOVA with two within-subjects factors: conflict type (low and high) and drug (placebo, 1 mg lorazepam, 2 mg lorazepam). Decision time data (ms) were analyzed with repeated-measures ANOVA with two within-subjects factors: dilemma type (moral-personal, moral-impersonal, nonmoral) and drug (placebo, 1 mg lorazepam, 2 mg lorazepam). Separate repeated-measures ANOVAs and simple contrasts were then used to explore the specificity of drug effects within each response type.  

**Results**

Table 1 presents descriptive statistics for task variables (ruthlessness scores and decision times) by dilemma type and drug condition for the whole sample and for male and female participants separately. There were no significant differences between sexes on any variable ($p > .05$).  

**Drug Effects on Dilemma Response Choices**

Scores on ruthlessness differed significantly between moral-impersonal, moral-personal, and nonmoral dilemmas, $F(2, 38) = 286.8, p < .001, \eta^2_p = 0.880$. This very large main effect indicates that dilemma type strongly altered the ruthlessness of participants’ response choices, providing background support for the notion that the content of a dilemma can influence response preferences. There was also a significant interaction between drug and dilemma type, $F(4, 36) = 4.7, p = .001, \eta^2_p = 0.107$, but no significant main effect of drug, $F(2, 38) = 1.4, p = .260, \eta^2_p = 0.034$, suggesting that lorazepam differentially altered response choices, depending upon the dilemma content. To probe the nature of this difference, we conducted separate repeated-measures ANOVAs for each dilemma type, revealing that lorazepam significantly increased ruthlessness in moral-personal dilemmas, $F(2, 38) = 5.6, p = .005, \eta^2_p = 0.126$. Furthermore, a clear dose–response effect of loraz-
epam on ruthlessness in moral-personal dilemmas was indicated by simple contrasts: placebo versus 1 mg, $F(1, 39) = 5.9$, $p = .020$, $\eta^2_p = 0.131$; placebo versus 2 mg, $F(1, 39) = 11.8$, $p = .001$, $\eta^2_p = 0.232$ (see Figure 1). Lorazepam did not exert a statistically significant effect on response choices in moral-impersonal dilemmas, $F(2, 38) = 1.9$, $p = .157$, $\eta^2_p = 0.046$, nor on response choices in nonmoral dilemmas, $F(2, 38) = 2.5$, $p = .092$, $\eta^2_p = 0.059$, relative to placebo.

As a check on the robustness of the lorazepam result, we then repeated the analysis, this time including two variables that are known a priori to be related to anxiety: sex and personality (as measured by the Fear Survey Schedule [FSS]; Wolpe & Lang, 1969). Table 1

Table 1
Descriptive Statistics for Ruthlessness Scores and Decision Times (in Seconds) by Dilemma Type and Drug Condition for the Whole Sample and for Male and Female Participants Separately

<table>
<thead>
<tr>
<th>Variable</th>
<th>Whole sample mean (SD)</th>
<th>Male participants mean (SD)</th>
<th>Female participants mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ruthlessness (Nonmoral, placebo)</td>
<td>5.53 (0.68)</td>
<td>5.45 (0.69)</td>
<td>5.60 (0.68)</td>
</tr>
<tr>
<td>2. Ruthlessness (Nonmoral, 1 mg lorazepam)</td>
<td>5.30 (0.76)</td>
<td>5.30 (0.66)</td>
<td>5.30 (0.86)</td>
</tr>
<tr>
<td>3. Ruthlessness (Nonmoral, 2 mg lorazepam)</td>
<td>5.15 (1.03)</td>
<td>5.15 (1.04)</td>
<td>5.15 (1.04)</td>
</tr>
<tr>
<td>4. Ruthlessness (Moral-impersonal, placebo)</td>
<td>3.70 (0.76)</td>
<td>3.90 (0.72)</td>
<td>3.50 (0.76)</td>
</tr>
<tr>
<td>5. Ruthlessness (Moral-impersonal, 1 mg lorazepam)</td>
<td>3.95 (0.75)</td>
<td>4.05 (0.83)</td>
<td>3.85 (0.67)</td>
</tr>
<tr>
<td>6. Ruthlessness (Moral-impersonal, 2 mg lorazepam)</td>
<td>4.03 (0.97)</td>
<td>3.95 (1.19)</td>
<td>4.10 (0.72)</td>
</tr>
<tr>
<td>7. Ruthlessness (Moral-impersonal, placebo)</td>
<td>1.75 (0.95)</td>
<td>1.70 (1.03)</td>
<td>1.80 (0.89)</td>
</tr>
<tr>
<td>8. Ruthlessness (Moral-personal, 1 mg lorazepam)</td>
<td>2.13 (1.14)</td>
<td>2.05 (1.15)</td>
<td>2.20 (1.15)</td>
</tr>
<tr>
<td>9. Ruthlessness (Moral-personal, 2 mg lorazepam)</td>
<td>2.33 (1.05)</td>
<td>2.30 (1.08)</td>
<td>2.35 (1.04)</td>
</tr>
<tr>
<td>10. Decision time (Nonmoral, placebo)</td>
<td>29.95 (8.25)</td>
<td>30.60 (6.16)</td>
<td>29.31 (10.05)</td>
</tr>
<tr>
<td>11. Decision time (Nonmoral, 1 mg lorazepam)</td>
<td>35.19 (10.86)</td>
<td>36.69 (9.30)</td>
<td>33.69 (12.29)</td>
</tr>
<tr>
<td>12. Decision time (Nonmoral, 2 mg lorazepam)</td>
<td>44.46 (16.70)</td>
<td>43.02 (14.41)</td>
<td>45.89 (19.00)</td>
</tr>
<tr>
<td>13. Decision time (Moral-impersonal, placebo)</td>
<td>22.22 (7.93)</td>
<td>22.99 (5.86)</td>
<td>21.44 (9.67)</td>
</tr>
<tr>
<td>14. Decision time (Moral-impersonal, 1 mg lorazepam)</td>
<td>27.83 (10.43)</td>
<td>29.18 (10.13)</td>
<td>26.47 (10.80)</td>
</tr>
<tr>
<td>15. Decision time (Moral-impersonal, 2 mg lorazepam)</td>
<td>34.03 (14.60)</td>
<td>35.13 (17.37)</td>
<td>32.93 (11.54)</td>
</tr>
<tr>
<td>17. Decision time (Moral-personal, 1 mg lorazepam)</td>
<td>33.30 (19.35)</td>
<td>31.26 (9.84)</td>
<td>35.32 (25.75)</td>
</tr>
<tr>
<td>18. Decision time (Moral-personal, 2 mg lorazepam)</td>
<td>38.00 (21.64)</td>
<td>38.93 (24.91)</td>
<td>37.07 (18.41)</td>
</tr>
</tbody>
</table>

Note. $N = 40$ (20 male). There were no significant differences in mean scores between sexes ($p < .05$).

Figure 1. Effect of lorazepam on dilemma responses. In the moral-personal but not the impersonal or nonmoral dilemmas, as the lorazepam dose increased, dilemma responses became significantly more ruthless in a linear manner: linear, $F(1, 39) = 11.8$, $p = .001$, $\eta^2_p = 0.232$; quadratic, $F(1, 39) = 0.3$, $p = .578$, $\eta^2_p = 0.008$. This dose-dependent drug effect was confirmed by simple contrasts: placebo versus 1 mg, $F(1, 39) = 5.9$, $p = .020$, $\eta^2_p = 0.131$; placebo versus 2 mg, $F(1, 39) = 11.8$, $p = .001$, $\eta^2_p = 0.232$. Error bars indicate one standard error of the mean. * $p < .05$. ** $p < .01$. 

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
This step was required to assess the possibility that the apparent effect of lorazepam on moral judgment was an artifact of preexisting sex or personality differences in anxiety proneness. The analysis of covariance revealed that when FSS scores and participant sex were included in the analysis, the main effect of dilemma and the drug × dilemma interaction remained significant, $F(2, 38) = 67.4, p < .001, \eta^2_p = 0.645$; $F(4, 36) = 3.5, p = .010, \eta^2_p = 0.252$. However, there was no significant interaction effect of sex or FSS scores, in any permutation of two- or three-way interactions ($p > .05$). This pattern of results suggests that the lorazepam result is not an artifact of preexisting differences in sex or personality and may instead reflect the hypothesized capacity of lorazepam to reduce anxiety elicited by moral-personal dilemmas.

**Drug Effects on Response Choices to Low-Conflict and High-Conflict Moral-Personal Dilemmas**

There was a very large main effect on moral personal dilemma response choices of conflict type (low and high), $F(1, 39) = 294.3, p < .001, \eta^2_p = 0.883$, as well as a strongly significant main effect of drug (placebo, lorazepam 1 mg, lorazepam 2 mg), $F(2, 38) = 7.6, p = .001, \eta^2_p = 0.163$. However, there was no significant interaction between conflict type and drug, $F(2, 38) = 0.2, p = .814, \eta^2_p = 0.005$, indicating that lorazepam effects on response choices did not differ significantly between low- and high-conflict moral-personal dilemmas.

**Drug Effects on Dilemma Decision Times**

Because lorazepam causes sedation in addition to reducing anxiety (e.g., Van Ruitenbeek et al., 2010), a demonstration that the lorazepam effect on response choices in moral personal dilemmas was not a side effect of sedation was required. Table 1 shows that decision times contained a large degree of variation, especially as the lorazepam dose increased. This means that any results pertaining to decision times are tentative. Nevertheless, decision time was significantly affected by drug condition, $F(2, 36) = 28.9, p < .001, \eta^2_p = 0.432$, and dilemma type, $F(2, 36) = 12.8, p < .001, \eta^2_p = 0.252$, but there was no significant interaction between drug and dilemma type, $F(4, 34) = 0.5, p = .732, \eta^2_p = 0.013$. The large main effect of drug condition indicated that lorazepam caused a linear increase in decision time in all three dilemma types. As a further check on this result we determined whether there was a significant association between the change in response times and change in ruthlessness scores for placebo versus 2 mg lorazepam for moral-personal dilemmas. There was no significant association (Pearson’s $r = -0.06, p = .715$), indicating that the increase in decision time caused by lorazepam was not responsible for biasing moral judgment toward ruthlessness in moral-personal dilemmas.

**Discussion**

Neuroimaging research suggests that moral dilemmas that involve directly harming other humans elicit stronger emotional reactions than either moral dilemmas where other humans are indirectly harmed or dilemmas in which there is no moral component (Greene et al., 2001). These seemingly automatic emotional reactions elicited by the prospect of directly harming others are thought to be a reflexive evolutionary adaptation that makes us averse to harming others directly and that must be overcome by cognitive effort in order for directly harmful acts to be committed (Greene et al., 2004). On the basis of clinical observations that interpersonal situations are especially potent eliciting stimuli for anxiety (American Psychiatric Association, 1994), we had hypothesized that anxiety is likely to be an emotion elicited in normal humans by the prospect of having to harm other people directly and is therefore likely to be a key factor in inhibiting behavior that harms others. We tested this idea by administering the anti-anxiety drug lorazepam (File, Bond, & Lister, 1982; Gould et al., 1997) to participants in a moral dilemma task. We predicted that it would dampen inhibitions concerning the commission of directly harmful acts in moral-personal dilemmas, thus increasing participants’ willingness to endorse ruthless acts.

In line with our prediction, we found that lorazepam significantly altered response choices in moral-personal dilemmas but not in moral-impersonal dilemmas or nonmoral dilemmas. The specific direction of the effect of lorazepam in moral-personal dilemmas was to cause a dose-dependent increase in participants’ preference in moral-personal dilemmas for responses that directly harm other humans. Our results are congruent with previous neuroimaging findings that emotional brain circuits are preferentially engaged by moral-personal dilemmas (Greene et al., 2001, 2004), but they extend this earlier work by suggesting that anxiety may be the specific emotion elicited by interpersonal moral dilemmas.

As a control, we tested the alternative possibility that lorazepam’s effects on moral judgment were caused by its sedative side effects (Van Ruitenbeek et al., 2010). We found that lorazepam slowed decision time equally across all three dilemma types and the magnitude of this change was not associated with the lorazepam-induced increase in preference for harmful response choices during the moral-personal dilemmas. This pattern of results argues against a sedative explanation of our dilemma decision finding.

Finally, we contrasted effects of lorazepam on low-conflict and high-conflict moral-personal dilemmas (Greene et al., 2008; Koenigs et al., 2007). These two types of moral personal dilemmas differ in that the high-conflict dilemmas test participants’ willingness to commit directly harmful acts for the greater good (i.e., for utilitarian reasons), whereas the low-conflict dilemmas test participants’ willingness to commit directly harmful acts for less noble (i.e., selfish) reasons. We found that lorazepam shifted the moral judgment of our participants toward endorsing directly harmful acts in low-conflict and high-conflict moral-personal dilemmas. Because both types of moral-personal dilemmas index willingness to cause direct harm to other people but only the high-conflict dilemmas index utilitarianism in the true philosophical sense, this result suggests that lorazepam is affecting variance in dilemma responses that is related to general willingness to cause harm rather than variance specifically related to utilitarianism (i.e., lorazepam makes us more willing to harm people directly whether or not that harm is for the greater good).

**Philosophical Implications**

We would argue, based on our results, that lorazepam increases ruthlessness, raising the interesting philosophical questions con-
This idea is congruent with clinical observations that psychopaths are underpinned by low activity in anxiety-generating emotion (Bartels & Pizarro, 2011), it might be suggested that psychopathy are also more utilitarian than average in their moral judgments. As individuals with high Implications for Psychopathy Research.

One interesting question arising from our findings is their relevance to research on psychopathy. As individuals with high scores on questionnaire measures of psychopathy-spectrum traits are also more utilitarian than average in their moral judgments (Bartels & Pizarro, 2011), it might be suggested that psychopathy is underpinned by low activity in anxiety-generating emotion systems that usually prompt aversion to harming others directly. This idea is congruent with clinical observations that psychopaths tend to be unusually free from anxiety (Cleckley, 1988) and neuroimaging studies suggesting that, relative to controls, psychopaths respond to emotional cues with significantly less activity in affective brain regions, such as the anterior and posterior cingulate, inferior frontal gyrus, amygdala/hippocampal formation, and ventral striatum (e.g., Kiehl, Smith, Hare, Forster, & Liddle, 2001). However, the existence of low-anxiety and high-anxiety psychopaths (e.g., Lykken, 1957; Newman, Patterson, Howland, & Nichols, 1990) and the finding that the former are more utilitarian than the latter in moral-personal dilemmas, but that both types of psychopaths were more utilitarian than healthy controls in moral-impersonal dilemmas (Koenigs et al., 2012), suggests a more nuanced interpretation is required. Integrating these earlier findings with our results, we tentatively suggest that a cognitive style that places little value on the fair treatment of others is the true root of psychopathy, but that this cognitive style is more freely expressed in interpersonal situations in the absence of anxiety, an expression that has the appearance (if not the philosophical basis) of utilitarianism.

An intriguing future study would therefore be to test the effect of anxiogenic interventions, such as cholecystokinin, yohimbine hydrochloride, or CO2 inhalation (Bailey, Kendrick, Diaper, Potokar, & Nutt, 2007), on the dilemma responses of healthy individuals and diagnosed psychopaths. If such interventions made psychopaths less likely to endorse responses that entail direct harm to other people, it could be accepted with some certainty that expression of their anti-social traits was influenced by anxiety-related harm inhibition. Studies of this type, if successful, would suggest that individuals diagnosed with psychopathy, once viewed as untreatable, might have their socially harmful behavior muted or even alleviated by anxiogenic drugs. Additionally, the proven efficacy of cognitive behavior therapy (CBT) in reducing anxiety (Gould et al., 1997) means that some form of reverse CBT or attention bias modification toward threat might ultimately be used to increase anxiety in psychopaths and thereby further reduce their problematic behavior in interpersonal settings.

Methodological Limitations

The strength of our conclusions regarding the role of anxiety in psychopathy is limited by a number of factors, the most important of which is that we tested only healthy participants. A further limitation of the present study concerns the question of drug specificity and mechanisms. As we examined only one specific compound of one class of anti-anxiety drugs, we do not at present know whether and how other anti-anxiety compounds might affect decision making in this paradigm. It would therefore be desirable to test the effects of other compounds shown to influence interpersonal and emotional processes, such as oxytocin (e.g., Meyer-Lindenberg, Domes, Kirsch, & Heinrichs, 2011). Should such studies show that the effects on moral judgment of anti-anxiety drugs are general, regulatory authorities that license anti-anxiety medications and physicians who prescribe them should give thought to the side effects of such drugs on moral judgment even at a low dose.

Further conceptual drawbacks of the present research concern the possibility that the capacity of lorazepam to drive interpersonal moral judgment in a ruthless direction is not caused specifically by its anti-anxiety properties but by its action on some other, as yet
undefined, emotion. Additionally, there is evidence that chronic dosing with benzodiazepines may actually boost negative emotions (e.g., Garcia et al., 2000). Future work may therefore be required to tease apart different emotional effects relating to acute versus chronic benzodiazepine dosing strategies. As a counter to these possibilities, evidence from studies of drug effects on threat responses in rodents and humans suggests that anti-anxiety drugs specifically dampen negative emotion elicited by cues of threat (e.g., Blanchard et al., 2003; Patrick et al., 1996; Perkins et al., 2009; Riba et al., 2001; Sartory et al., 1990). The precise label for such an emotion is a matter of debate owing to the long-standing confusion concerning the semantic differentiation of anxiety from other closely related negative emotions such as fear (e.g., Geer, 1965). Nevertheless, as Koenigs et al. (2012) have already shown that individual differences in anxiety-proneness modulate the utilitarianism of psychopaths faced with interpersonal moral dilemmas, it would seem plausible that it is the anti-anxiety effects of lorazepam that are doing the work in our study, especially as sedation effects have been controlled for.

Conclusions

Bearing these limitations in mind, we conclude that lorazepam makes people more ruthless in general rather than boosting utilitarianism specifically but that the precise explanation for this effect requires much further research before it is fully elucidated. We favor the hypothesis that lorazepam increases ruthlessness by reducing perceptions of threat intensity that in turn dampen anxiety caused by the prospect of harming others, freeing participants to endorse harmful acts. We suggest this is the case because of our recent finding that lorazepam reduced the intensity of threat-avoidance behavior in healthy human participants (Perkins et al., 2009). That study had very different demand characteristics than the present investigation, yet the lorazepam effects are congruent: In both studies lorazepam reduced the avoidance of direct physical harm, whether to the person (in our 2009 study) or to other people (in the present study). This line of reasoning ultimately leads to an extremely interesting question: Why should the prospect of causing direct harm to others have evolved to be perceived in a way similar to a threat to oneself?

References

Kant, I. (1959). Foundation of the metaphysics of morals (L. W. Beck, Trans.). Indianapolis, IN: Bobbs-Merrill. (Original work published 1785)
Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007, April 19). Damage to the prefrontal cortex...


Vermeeren, A., Jackson, J. L., Muntjewerff, N. D., Quint, P. J., Harrison, E. M., & O’Hanlon, J. F. (1995). Comparison of acute alprazolam (0.25, 0.50 and 1.0 mg) effects versus those of lorazepam 2 mg and placebo on memory in healthy volunteers using laboratory and telephone tests. *Psychopharmacology, 118*, 1–9. doi:10.1007/BF02245243


Received February 19, 2012
Revision received August 14, 2012
Accepted August 16, 2012

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.