

# Serotonin selectively influences moral judgment and behavior through effects on harm aversion

Molly J. Crockett<sup>a,1</sup>, Luke Clark<sup>a</sup>, Marc D. Hauser<sup>b</sup>, and Trevor W. Robbins<sup>a</sup>

<sup>a</sup>Behavioural and Clinical Neuroscience Institute, Department of Experimental Psychology, University of Cambridge, Cambridge CB2 3EB, United Kingdom; and <sup>b</sup>Department of Psychology and Human Evolutionary Biology, Harvard University, Cambridge, MA 02138

Edited by Mortimer Mishkin, National Institute for Mental Health, Bethesda, MD, and approved August 9, 2010 (received for review July 1, 2010)

**Aversive emotional reactions to real or imagined social harms infuse moral judgment and motivate prosocial behavior. Here, we show that the neurotransmitter serotonin directly alters both moral judgment and behavior through increasing subjects' aversion to personally harming others. We enhanced serotonin in healthy volunteers with citalopram (a selective serotonin reuptake inhibitor) and contrasted its effects with both a pharmacological control treatment and a placebo on tests of moral judgment and behavior. We measured the drugs' effects on moral judgment in a set of moral 'dilemmas' pitting utilitarian outcomes (e.g., saving five lives) against highly aversive harmful actions (e.g., killing an innocent person). Enhancing serotonin made subjects more likely to judge harmful actions as forbidden, but only in cases where harms were emotionally salient. This harm-avoidant bias after citalopram was also evident in behavior during the ultimatum game, in which subjects decide to accept or reject fair or unfair monetary offers from another player. Rejecting unfair offers enforces a fairness norm but also harms the other player financially. Enhancing serotonin made subjects less likely to reject unfair offers. Furthermore, the prosocial effects of citalopram varied as a function of trait empathy. Individuals high in trait empathy showed stronger effects of citalopram on moral judgment and behavior than individuals low in trait empathy. Together, these findings provide unique evidence that serotonin could promote prosocial behavior by enhancing harm aversion, a prosocial sentiment that directly affects both moral judgment and moral behavior.**

morality | punishment | ultimatum game | empathy | emotion

Considerations of harm and care are central to human morality (1). Growing evidence supports the notion that empathic responses to the suffering of others are critical for motivating prosocial behavior (2, 3). This claim is perhaps most strikingly supported by the case of psychopathy, in which marked deficits in empathy and guilt are in large measure responsible for morally inappropriate behavior (4, 5). Emotional reactions to moral transgressions also appear to infuse moral judgment (6). People often judge harms to innocent victims as morally forbidden even when such harms potentially achieve superordinate goals, such as saving the lives of many others (7, 8). Further, such moral condemnation is especially strong when harms are emotionally salient (7, 9, 10) and when negative emotions are primed (11).

Recent work in psychology and neuroscience suggests that certain "prosocial moral sentiments" (12), including aversive emotional reactions to real or imagined harmful acts, are highly variable between individuals and across situations (13, 14). However, the mechanisms that drive this variability are not well understood. Here, we use a unique experimental approach to explore the causal role of neurochemical systems in dynamically shaping moral judgment and behavior. Specifically, we show that the neurotransmitter serotonin directly modifies subjects' moral judgments and behavior by means of enhancing aversion to personally harming others.

Serotonin is richly involved in the biology of social behavior across species, from the swarming of locusts (15) to the social norms of *Homo sapiens* (16). The serotonin system densely innervates structures previously implicated in moral judgment and behavior,

including the ventromedial prefrontal cortex (vmPFC), insula, and amygdala (7, 10, 17–21). Decades of research have shown that prosocial and affiliative behaviors are associated with intact or enhanced serotonin function, whereas antisocial and aggressive behaviors are associated with impaired or reduced serotonin function (22, 23).

A common explanation for the relationship between serotonin and prosocial behavior is that serotonin promotes the effortful control of violent impulses or down-regulation of emotional reactions to provocation (24, 25). Support for this emotion regulation account of serotonin in moral behavior is mainly based on observations that strong emotions often accompany the violent behavior that occurs in populations with impaired serotonin function (24, 25), and that brain regions implicated in representing and regulating emotions, such as the orbitofrontal cortex (26, 27), receive a high density of serotonergic input (17, 18).

However, antisocial behavior may also result from impaired aversive responses to the distress of others (5, 28). Such responses engage the amygdala and vmPFC (5), which also receive serotonergic inputs (18). Recent work highlighting serotonin's involvement in enhancing expectations of aversive outcomes (29, 30) suggests an alternative, harm aversion account of serotonin in prosocial behavior: that serotonin amplifies the aversiveness of personally harming others and, in so doing, promotes prosocial behavior while discouraging antisocial behavior. If this hypothesis is correct, then enhancing serotonin should cause considerations of harm to loom large in moral judgment and behavior, even if there are contravening utilitarian benefits or fairness goals.

This study was designed to test between the emotion regulation and harm aversion hypotheses, while providing a unique approach to understanding the causal role of the brain's neurotransmitters in moral judgment and behavior. We manipulated serotonin in healthy volunteers by using the highly selective serotonin reuptake inhibitor (SSRI) citalopram, which boosts serotonin neurotransmission by blocking its reuptake and prolonging its actions in the synapse (31). To probe the neurochemical specificity of serotonin in modulating moral judgment and behavior, we contrasted the effects of citalopram with those of atomoxetine, a relatively selective noradrenaline reuptake inhibitor that enhances noradrenaline neurotransmission (31). Because noradrenaline has been strongly linked to executive functions (32, 33), comparing the effects of citalopram with those of atomoxetine on moral judgment and behavior is relevant to the ongoing debate surrounding the role of executive control versus that of emotion in these processes (12, 34, 35).

Author contributions: M.J.C., L.C., M.D.H., and T.W.R. designed research; M.J.C. performed research; M.J.C. analyzed data; and M.J.C., L.C., M.D.H., and T.W.R. wrote the paper.

Conflict of interest statement: T.W.R. consults for Cambridge Cognition, Pfizer, GlaxoSmithKline, E. Lilly Inc., Lundbeck, Roche, Allon Therapeutics, and Pangenics.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

<sup>1</sup>To whom correspondence should be addressed. E-mail: mc536@cam.ac.uk.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1009396107/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1009396107/-DCSupplemental).

To examine the neurochemical modulation of moral judgment, we tested the effects of citalopram and atomoxetine on judgments in a series of moral dilemmas pitting utilitarian outcomes (e.g., saving five lives) against highly aversive harmful actions (e.g., killing an innocent person) that varied in emotional salience (10, 36, 37). Personal scenarios involve emotionally salient violent acts (e.g., pushing someone in front of a train to prevent it from hitting five people); previous work has shown that these scenarios activate brain regions implicated in emotional processing, including the vmPFC and amygdala (7, 36). In contrast, less emotionally salient impersonal scenarios (e.g., flipping a switch to divert a train to hit one person instead of five people) activate regions implicated in executive control, including lateral prefrontal and posterior parietal cortex (7, 36). Personal harms are generally judged to be less acceptable than impersonal harms leading to identical outcomes, presumably because the emotional aversion to “up close and personal” violence influences judgment (6, 7). Demonstrating a selective effect of citalopram on judgment in personal moral scenarios would therefore provide direct support for our hypothesis that serotonin modulates emotional aversion to harm, rather than other aspects of moral judgment.

Importantly, the emotion regulation and harm aversion accounts of serotonin function make opposite predictions about the direction of the effect of citalopram on moral judgment. If serotonin inhibits prepotent emotional responses, we might expect citalopram to blunt this emotional spur to moral judgment, thus increasing the acceptability of harming one person to save others in personal moral scenarios. In contrast, if serotonin amplifies the aversiveness of harming others, we would expect the opposite: citalopram should reduce the acceptability of harming one person to save others in personal moral scenarios.

We investigated the neurochemical modulation of moral behavior by using the ultimatum game (UG) (38). In the UG, one player (the responder) decides whether to accept or reject monetary offers from another player (the proposer). If the responder accepts, both players are paid accordingly; if he rejects, neither is paid. Responders tend to punish proposers who offer less than  $\approx 20\text{--}30\%$  of the shared stake by rejecting their “unfair” offers (39). Rejecting unfair offers enforces a fairness norm but also harms the other player financially; responders are thus forced to tradeoff concerns for fairness versus harm when considering unfair offers. We recently reported that temporarily lowering serotonin in healthy volunteers playing the role of responder increased their tendency to reject unfair offers in the UG (16). In this study, we sought to extend these findings by testing whether serotonergic modulation of behavior in the UG is bidirectional and neurochemically specific. As in our previous study, participants played the role of responder in a series of one-shot UGs, deciding whether to accept or reject offers ranging from relatively fair (45% of the stake) to unfair (20% of the stake). We predicted that augmenting serotonin neurotransmission with citalopram would have the opposite effect of serotonin depletion, decreasing responders’ rejection of unfair offers.

Note that such a result in the UG would be consistent with either proposed account of serotonin in prosocial behavior. Because several studies have shown that emotional reactions to unfairness motivate rejection (40–43), reduced rejection of unfair offers after citalopram could reflect enhanced regulation of emotional reactions (but see also ref. 44). Alternatively, decreased rejection of unfair offers could reflect increased aversion to harming the proposers. We were therefore critically interested in juxtaposing citalopram’s effects on UG behavior with its effects on moral judgment in the same volunteers. Observing that citalopram reduces rejection rates in the UG and increases the acceptability of harming one person to save many others would support a role for serotonin in enhancing self-control in social contexts. This set of findings would be consistent with

dual-process accounts of moral judgment and behavior, in which emotion and rational thought compete for behavioral output (34, 36). In contrast, observing that citalopram reduces the acceptability of harming one person to save many others in conjunction with reducing rejection rates in the UG would support a role for serotonin in enhancing aversive emotional reactions to the prospect of personally harming others. This set of findings would be consistent with a more integrated view of emotional and cognitive mechanisms in generating prosocial sentiments, which guide moral judgment and behavior (12).

Aversive emotional reactions to choices that harm others should be stronger in people high in empathy, who have a greater ability to share the affective experiences of others (3). Empathy goes beyond mere emotional mimicry; it is grounded in an observer’s perspective and motivates prosocial behavior (13). We therefore investigated how trait empathy and state serotonin interact in modulating moral judgment and behavior. Previous work has shown that serotonin manipulations often have stronger effects on behavior in individuals with high trait levels of the measured behavior; state by trait interactions have been demonstrated for aggression, impulsivity, and social dominance (45–47). In this study, we hypothesized that the effects of serotonin manipulations on moral judgment and behavior might be stronger in individuals high in trait empathy. We measured trait empathy by using the Interpersonal Reactivity Index (48), which predicts prosocial helping behavior (2) and modulates neural responses to other people’s pain (49). Demonstrating stronger prosocial effects of citalopram in highly empathic individuals would lend further support to our hypothesis that serotonin promotes prosocial behavior by means of enhancing harm aversion.

Healthy volunteers attended three sessions and received clinically relevant doses of citalopram (30 mg) and atomoxetine (60 mg) as well as placebo in a double-blind fully counterbalanced design. On each session, they completed the moral judgment task, the UG, and standard measures of executive function and mood.

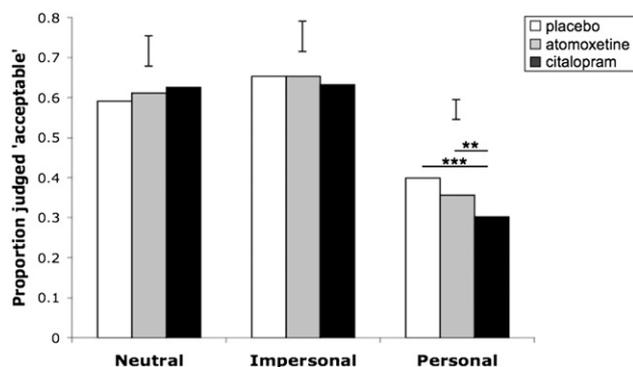
## Results

Consistent with previous findings, emotional salience influenced moral judgment. We compared participants’ judgments of the acceptability of harmful actions in three types of scenarios (nonmoral control scenarios, emotionally salient personal scenarios, and less emotionally salient impersonal scenarios) after administration of citalopram, atomoxetine, and placebo. We found a main effect of scenario type ( $\chi^2 = 42.990$ ,  $P < 0.001$ ); personal harms were judged to be less acceptable than impersonal harms ( $P < 0.001$ ) and nonmoral actions ( $P < 0.001$ ).

This pattern of moral judgment varied as a function of drug treatment, as indicated by a significant interaction between drug and scenario type ( $\chi^2 = 10.672$ ,  $P = 0.03$ ; Fig. 1). Post hoc pairwise comparisons showed no significant differences between drug treatments for judgments of neutral scenarios (all  $P > 0.238$ ) or impersonal scenarios (all  $P > 0.545$ ). However, in the emotionally salient personal scenarios, citalopram made subjects far more likely to judge harmful actions as forbidden, reducing acceptability judgments relative to both placebo ( $P = 0.001$ ) and atomoxetine ( $P = 0.048$ ); atomoxetine and placebo did not differ from one another ( $P = 0.250$ ).

We next examined the drugs’ influence on moral behavior in the UG as participants played the role of responder, by comparing participants’ rejection rates of fair (45%), unfair (30%), and most unfair (20%) offers after citalopram, atomoxetine, and placebo. As predicted, rejection rates increased with decreasing offer fairness (main effect of fairness:  $\chi^2 = 96.831$ ,  $P < 0.001$ ); fair offers were rejected less frequently than moderately unfair offers, which were rejected less frequently than the most unfair offers (all pairwise  $P < 0.001$ ).

The harm-avoidant bias in moral judgment observed after citalopram was also evident in behavior during the UG. Only the



**Fig. 1.** Effects of drug and scenario type on moral judgment. Citalopram reduced the acceptability of harms in emotionally salient personal scenarios, relative to both placebo and atomoxetine.  $**P \leq 0.05$ ;  $***P \leq 0.01$ . Error bars represent twice the SE of the difference of means (*SI Materials and Methods*).

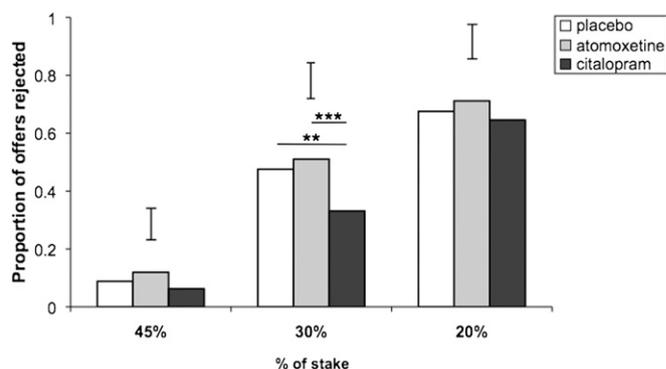
serotonin manipulation influenced responder behavior, as indicated by a significant drug  $\times$  fairness interaction ( $\chi^2 = 15.669$ ,  $P = 0.003$ ; Fig. 2). Citalopram reduced rejection of the unfair (30%) offers, relative to both atomoxetine ( $P = 0.009$ ) and placebo ( $P = 0.034$ ), which did not differ ( $P = 0.520$ ).

After completing the UG, participants rated the fairness of a representative sample of offers received in the UG, but neither drug treatment influenced judgments of fairness. There was a significant main effect of fairness category on fairness ratings of offers ( $P < 0.001$ ); as expected, after all three treatments, 45% offers were rated as more fair than 30% offers and 20% offers (all pairwise  $P < 0.001$ ), but there were no significant effects of drug or session (all  $P > 0.12$ ) on fairness ratings.

We did not observe significant effects of the noradrenaline-enhancing drug atomoxetine on moral judgment or UG behavior, but atomoxetine did enhance performance on standard measures of executive control (*SI Results*), consistent with previous findings (32). Meanwhile, citalopram had no effect on executive control. Neither drug significantly influenced self-reported mood, and although both citalopram and atomoxetine produced mild nausea relative to placebo, physical side effects are unlikely to be responsible for the effects of citalopram on moral judgment and behavior (*SI Results*).

Next, we tested the hypothesis that the prosocial effects of citalopram on moral judgment and behavior would be stronger in individuals high in trait empathy. We examined the moderating influence of trait empathy on the neuromodulation of moral judgment by testing the effects of drug and scenario type on moral judgment, including trait empathy as a covariate. This analysis revealed main effects of scenario type ( $\chi^2 = 8.557$ ,  $P = 0.014$ ) and drug ( $\chi^2 = 7.425$ ,  $P = 0.024$ ), a significant interaction between drug and trait empathy ( $\chi^2 = 10.811$ ,  $P = 0.004$ ), and a significant interaction between drug, scenario type, and trait empathy ( $\chi^2 = 11.557$ ,  $P = 0.021$ ). Because there were more females than males in the high empathy group, we included gender as a separate variable in this analysis; however, gender did not significantly impact moral judgment, and it did not interact significantly with any other variable.

To explore the interactions, we performed a median split on the empathy scores, dividing subjects into high and low empathy groups, and repeated the above analysis separately in each group. In the low empathy group, there was a significant main effect of scenario type ( $\chi^2 = 14.949$ ,  $P = 0.001$ ), but the drug  $\times$  scenario type interaction was not significant ( $\chi^2 = 3.372$ ,  $P = 0.498$ ). In the high empathy group, there was a main effect of scenario type ( $\chi^2 = 32.950$ ,  $P < 0.001$ ) and a significant drug  $\times$  scenario type interaction ( $\chi^2 = 14.020$ ,  $P = 0.007$ ). Post hoc comparisons confirmed these findings: In the low empathy group, there were no



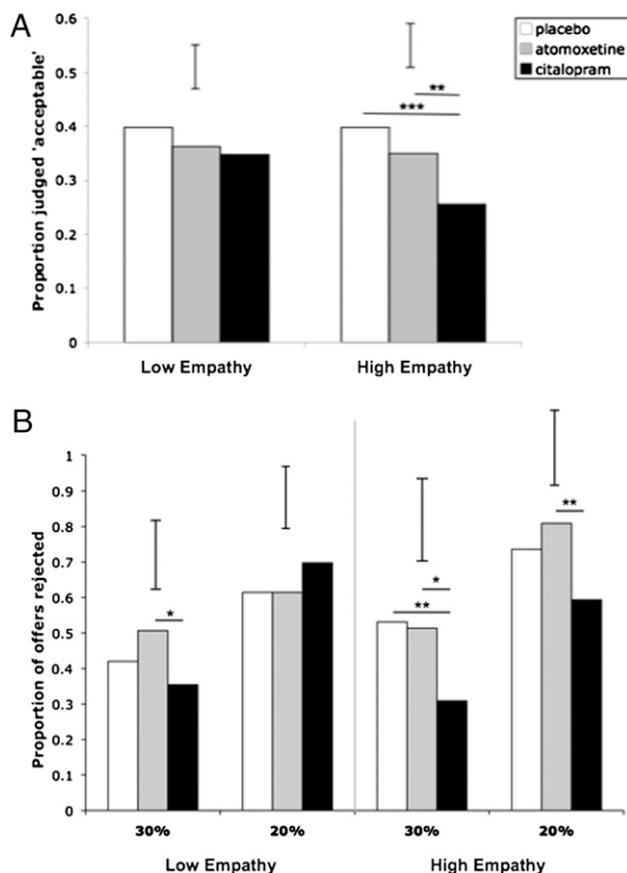
**Fig. 2.** Effects of drug and offer fairness on rejection rates in the UG. Citalopram reduced rejection rates of unfair offers, relative to both atomoxetine and placebo.  $**P \leq 0.05$ ;  $***P \leq 0.01$ . Error bars represent twice the SE of the difference of means.

significant differences between drug treatments for any scenario type (all  $P > 0.209$ ). In the high empathy group, however, citalopram significantly reduced acceptability judgments of personal harms relative to both placebo ( $P < 0.001$ ) and atomoxetine ( $P = 0.011$ ). Thus, it appears that the effects of citalopram on judgment of personal harms observed in the whole group are driven by subjects high in trait empathy (Fig. 3A).

To test whether trait empathy moderated the drug effects on behavior in the UG, we analyzed the effects of drug and fairness on rejection rates, including trait empathy as a covariate. This analysis revealed a main effect of offer fairness ( $\chi^2 = 98.959$ ,  $P < 0.001$ ), a drug  $\times$  fairness interaction ( $\chi^2 = 15.750$ ,  $P = 0.003$ ), and a three-way interaction between drug, fairness, and trait empathy ( $\chi^2 = 10.814$ ,  $P = 0.029$ ). Gender did not affect rejection rates, and it did not interact significantly with any other variable. As with the moral judgment analysis, we split our sample into a high and a low empathy group based on their trait empathy scores and performed separate analyses in each group. The main effect of fairness and the drug  $\times$  fairness interaction were significant in both the high and low empathy groups, but post hoc comparisons revealed different drug effects on rejection rates as a function of trait empathy (Fig. 3B). In the low empathy group, citalopram showed a trend toward reducing rejections of unfair (30%) offers, relative to atomoxetine ( $P = 0.063$ ) but not placebo ( $P = 0.341$ ); no other pairwise comparisons reached significance in this group. In contrast, the effects of citalopram were stronger in the high empathy group. Citalopram showed a trend toward reducing rejection of unfair (30%) offers relative to atomoxetine ( $P = 0.060$ ), and significantly reduced rejection rates relative to placebo ( $P = 0.05$ ); in the high empathy group, citalopram also reduced rejection of the most unfair (20%) offers, relative to atomoxetine ( $P = 0.027$ ). As with moral judgment, the prosocial effects of citalopram on UG behavior appear to be driven by subjects high in trait empathy.

## Discussion

The goal of this study was to examine the modulatory role of serotonin on human moral judgment and behavior. We were specifically interested in determining whether serotonin promotes prosocial behavior by supporting regulation of prepotent emotional reactions, or by enhancing the aversiveness of harming others. Our results support the harm aversion account of serotonin in prosocial behavior. Enhancing serotonin function with citalopram selectively influenced moral judgment in the emotionally salient personal scenarios; relative to placebo and the positive control atomoxetine, citalopram made subjects less likely to judge personal harms as permissible, without affecting judgments of impersonal harms. The fact that citalopram only influenced judgment in the emotionally



**Fig. 3.** Neuromodulation of moral judgment and behavior as a function of trait empathy. (A) In the emotionally salient personal scenarios, the effects of citalopram on moral judgment were stronger in individuals high in trait empathy. (B) The effects of citalopram on rejection of unfair (20% and 30%) offers were stronger in individuals high in trait empathy. \* $P \leq 0.10$ , \*\* $P \leq 0.05$ ; \*\*\* $P \leq 0.01$ . Error bars represent twice the SE of the difference of means.

salient personal scenarios, which evoke strong emotional reactions relative to impersonal scenarios (7, 10), supports our hypothesis that serotonin enhances aversive emotional reactions to harm, which fits with serotonin's purported role in the neurobiology of punishments and threats (29, 30).

A citalopram-induced harm-avoidant bias was also evident in behavior during the UG. Complementing our previous finding that serotonin depletion increased rejection of unfair offers in the UG (16), we found that citalopram affected UG behavior in the opposite direction, reducing rejection of unfair offers. The nature of the motivation underlying punitive rejection behavior in the UG is ambiguous. Rejecting unfair offers is simultaneously prosocial at the group level, because it enforces fairness norms (50), but antisocial at the individual level, because it harms the proposer (22). The present findings imply that serotonin promotes prosocial behavior at the individual level; citalopram reduced both the willingness to endorse harming another person in hypothetical scenarios, and the willingness to harm another person in a real economic transaction. This set of findings fits with the extensive literature implicating reduced or impaired serotonin function in aggression (23) and supports our hypothesis that serotonin mitigates aggressive impulses by amplifying the aversiveness of harming others, rather than by improving emotion regulation. This effect was reflected in distinct measures of both judgment and behavior.

Notably, our pattern of results mirrors that observed in patients with damage to the vmPFC. After citalopram, our subjects were less likely to endorse personal moral harms and less likely to punish unfairness in the UG; previous studies have shown that vmPFC patients are more likely to endorse personal moral harms (10, 51), and more likely to punish unfairness in the UG (19, 52). One explanation of such results is that vmPFC patients suffer from a selective deficit in prosocial sentiments such as guilt and empathy (12, 53). According to this account, our findings imply that serotonin enhances prosocial sentiments: After citalopram, subjects were less likely to advocate harming an innocent bystander and more likely to "turn the other cheek" and forgive unfair behavior. Whether these effects of manipulating serotonin arise specifically from its modulatory influence on vmPFC (or associated regions such as the insula) cannot be determined from the present findings, but this is a plausible hypothesis that could be tested further by using functional neuroimaging. Indeed, functional MRI studies have shown that the vmPFC processes prosocial moral sentiments, such as guilt and compassion, as distinct from other-critical feelings such as anger and indignation, processed in lateral prefrontal regions (12, 54). Future studies should address how serotonin modulates the balance between anger, which could drive altruistic punishment, versus guilt and compassion, which could restrain punishment (12).

The prosocial effects of citalopram on moral judgment and behavior in the UG were stronger in individuals high in trait empathy. This finding lends further support to the hypothesis that serotonin enhances aversive reactions to social harms, because these reactions are stronger in individuals high in trait empathy (3). We suggest that serotonin modulates empathic responses to harm—i.e., by boosting an already-present neural signal—rather than being the source of empathic responses. No studies have examined the influence of serotonin manipulations on more direct measures of empathy, such as neural responses to perceived pain in others, a promising avenue for future studies. Empathic neural responses to others' suffering involve the insula (3, 49), which receives a high density of serotonin modulation (17), and serotonin manipulations influence the detection of emotion in faces (55), which has been linked to trait empathy (56). Serotonin also promotes the release of oxytocin and vasopressin (57), neuropeptides strongly implicated in empathy and prosocial behavior (58).

The fact that the prosocial effects of citalopram were stronger in more highly empathic individuals also has consequences for the pharmacological treatment of aggressive and antisocial behavior. Serotonin has been implicated in a wide range of psychiatric disorders with antisocial and aggressive symptoms, including intermittent explosive disorder, antisocial personality disorder, and psychopathy (59–61), and these disorders are often treated with SSRIs such as citalopram (62). Our findings suggest that trait empathy measures could predict whether patients are likely to respond to SSRI treatment and imply that such treatments are less likely to succeed in psychopaths and patients with vmPFC damage, both of whom show a marked lack of empathy (63). A role for serotonin in prosocial sentiments may also have implications for understanding the excessive feelings of guilt in depressed patients (64).

Citalopram did not alter self-reported mood or executive control, indicating that the effects of citalopram on moral judgment and behavior cannot be explained by indirect effects on mood or cognitive function. However, manipulating noradrenaline with atomoxetine had a clear effect on executive function; in line with previous findings (32, 33), atomoxetine improved performance on two standard measures of cognitive control. But atomoxetine had no effect on moral judgment or behavior in the UG. Our findings therefore do not support the hypothesis that executive control processes compete with emotional reactions to harm in moral judgment, with the former promoting utilitarian

responses (7); if this were the case, then atomoxetine (which enhanced executive function) would have increased the acceptability of harming one to save many in personal moral scenarios. Our findings also do not support the hypothesis that executive control processes compete with selfish impulses in the UG, with the former promoting rejection of unfair offers (50); if this were the case, then atomoxetine would have increased rejection of unfair offers in the UG. Our findings instead highlight the primacy of prosocial sentiments involving considerations of harm in shaping moral judgment and social behavior.

A limitation of this study is that both citalopram and atomoxetine induced nausea in some of our participants, although there was no significant difference between the drugs' induced nausea. Previous studies have shown that disgust covaries with moral judgment (11) and punishment in the UG (43). However, it is unlikely that our pattern of results is simply due to drug-induced nausea, because only citalopram influenced moral judgment and UG behavior, even though nausea was slightly higher on atomoxetine; further, the effects of nausea on moral judgment went in the opposite direction from those of citalopram.

This study demonstrates that altering central serotonin function in healthy volunteers has selective causal effects on moral judgment and social behavior. Blocking serotonin reuptake with citalopram influenced moral judgment in emotionally salient personal scenarios, making subjects less likely to endorse harming one person to save many others, and also made subjects less likely to harm others via punishment in an economic game, effects that were stronger in highly empathic individuals. This pattern of results implies that serotonin promotes prosocial behavior by enhancing the aversiveness of harming others, an effect that drives both moral judgment and behavior. Our findings also have implications for the use of serotonin agents in the treatment of antisocial and aggressive behavior (47, 62). Understanding the influence of serotonin on social and moral behavior is especially important because serotonin is implicated in a wide range of psychiatric disorders and sensitive to the environmental context (33), which is a demonstrably powerful force shaping our social lives.

## Materials and Methods

**Participants.** The protocol was approved by the Cambridgeshire Research Ethics Committee. Thirty healthy subjects (13 males; mean age, 25.6 y) were screened for neurological and psychiatric disorders by a consultant psychiatrist; each subject gave written informed consent before participating in the study. Exclusion criteria were any history of cardiac, hepatic, renal, pulmonary, neurological, psychiatric or gastrointestinal disorders, medication/drug use, and personal or family history of major depression or bipolar affective disorder. Participants were financially compensated. Two participants dropped out of the study before completing all three sessions. Two participants were excluded from all analyses because of severe peripheral side effects, and two participants were excluded for indicating at debriefing that they did not believe the UG was real. The final analysis was carried out in 24 participants.

**General Procedure.** Participants attended three sessions at Addenbrooke's Hospital in Cambridge, UK (at least 1 wk apart) and received single doses of atomoxetine (60 mg), citalopram (30 mg), and placebo in a double-blind fully counterbalanced design. We selected doses that were clinically relevant according to established treatment guidelines for attention deficit hyperactivity disorder (atomoxetine) and obsessive-compulsive disorder and depression (citalopram; British National Formulary; [www.bnf.org](http://www.bnf.org)), and in line with a previous study in our laboratory (65). At the start of each session, participants completed mood and trait questionnaires and took the drug orally. Administration of cognitive testing was timed to coincide with the peak effects of both compounds, based on previous pharmacokinetic data. After spending 1.5 h in a quiet area, participants completed a mood questionnaire and the UG, moral judgment, and executive function tasks (*SI Materials and Methods*) as part of a larger cognitive testing battery. Mood was assessed by using visual analog

scales and the Positive and Negative Affect Scale (66). At the end of the third session, participants completed a debriefing questionnaire about their overall impressions of the study, including whether they believed they would be paid based on their choices during the UG and whether they had any suspicions about the order of drug administration.

**Moral Judgment Task.** The moral judgment task was adapted from previous studies of moral reasoning in healthy volunteers (36, 37) and brain lesion patients (10); all stimulus materials have been published elsewhere. In the task, participants made judgments on a series of hypothetical scenarios, presented as text on three screens. The first two screens described the scenario, and the third screen posed a question relevant to the current scenario (e.g., "Is it acceptable to...?"). Participants had unlimited time to read the scenarios and respond to the question, pressing the spacebar to advance through the scenario text and responding to the question with keys labelled "yes" and "no." Responding yes always indicated endorsement of the proposed action.

Participants responded to two types of moral scenarios designed to contrast the impact of emotional salience on moral judgment: (i) emotionally salient "personal" harms that involved harming one person to benefit many and (ii) "impersonal" harms that were less emotionally salient than the personal cases but also involved harming one person while benefiting many. Each subject also responded to a set of nonmoral scenarios. On each session, participants responded to 6 nonmoral scenarios, 6 impersonal moral scenarios, and 17 personal moral scenarios. The complete set of personal scenarios described both avoidable harms and inevitable harms (37), whereas the set of impersonal scenarios described only avoidable harms; therefore, to better match the personal and impersonal scenarios we excluded the personal scenarios describing inevitable harms ( $n = 8$ ) from the current analysis, focusing only on the personal scenarios describing avoidable harms ( $n = 9$ ). Participants completed a different set of matched scenarios on each session to avoid repetition effects; task order was fully counterbalanced across sessions independently from drug order counterbalancing. The dependent measure of interest was the proportion of scenarios in each category endorsed as "acceptable."

**UG Task.** We used the same UG task as in our first study (16). Participants played the role of responder via computer interface. On each session, participants played the role of responder in 24 games, each with a different proposer. Proposer identities were randomly matched with offers. There were eight fair offers, ranging from 40% to 50% of the stake; eight unfair offers, ranging from 27% to 33% of the stake; and eight extremely unfair offers, ranging from 18% to 22% of the stake, presented in random order. Participants received identical offers on each session. On different trials, the same monetary amount could appear as a large percentage of the total stake and therefore "fair," or as a small percentage of the total stake and therefore "unfair." This design controlled for any effects of offer magnitude. After completing the UG task, participants rated the fairness of six offers representative of the different fairness categories on a Likert scale of 1 (very unfair) to 7 (very fair). The critical dependent measures were the proportions of offers rejected at each level of fairness and the fairness ratings at each level of fairness.

**Data Analysis.** In line with previous studies in this area (10, 19, 52), response data were modeled by using generalized estimating equations (GEE) (67). For the UG analysis, we analyzed rejection rates with drug, session, and fairness as within-subjects factors. For the Moral Judgment analysis, we analyzed acceptable judgment rates with drug, session, and scenario type as within-subjects factors. Factors were dropped from subsequent analyses when nonsignificant. When significant factors were found, we conducted post hoc pairwise Least Significant Difference tests (for non-normally distributed dependent variables) and  $t$  tests (for normally distributed dependent variables). Significant differences were set at  $P < 0.05$ .

**ACKNOWLEDGMENTS.** We thank the staff at the Wellcome Trust Clinical Research Facility, B. Huebner, K. Craig, and U. Muller. This work was completed within the University of Cambridge Behavioural and Clinical Neuroscience Institute, funded by a joint award from the Medical Research Council and the Wellcome Trust and also by a Network Award from the JT McDonnell Foundation.

- Haidt J (2007) The new synthesis in moral psychology. *Science* 316:998–1002.
- Davis MH (1983) The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach. *J Pers* 51:167–184.

- Singer T, Lamm C (2009) The social neuroscience of empathy. *Ann N Y Acad Sci* 1156:81–96.
- Hare RD (2003) *The Hare Psychopathy Checklist—Revised* (Multi-Health Systems, Toronto), 2nd Ed.

5. Blair RJR (2007) The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends Cogn Sci* 11:387–392.
6. Haidt J (2001) The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychol Rev* 108:814–834.
7. Greene JD, Sommerville RB, Nystrom LE, Darley JM, Cohen JD (2001) An fMRI investigation of emotional engagement in moral judgment. *Science* 293:2105–2108.
8. Schaich Borg J, Hynes C, Van Horn J, Grafton S, Sinnott-Armstrong W (2006) Consequences, action, and intention as factors in moral judgments: An fMRI investigation. *J Cogn Neurosci* 18:803–817.
9. Greene JD, et al. (2009) Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition* 111:364–371.
10. Koenigs M, et al. (2007) Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* 446:908–911.
11. Schnall S, Haidt J, Clore GL, Jordan AH (2008) Disgust as embodied moral judgment. *Pers Soc Psychol Bull* 34:1096–1109.
12. Moll J, de Oliveira-Souza R (2007) Moral judgments, emotions and the utilitarian brain. *Trends Cogn Sci* 11:319–321.
13. de Vignemont F, Singer T (2006) The empathic brain: How, when and why? *Trends Cogn Sci* 10:435–441.
14. Hein G, Singer T (2008) I feel how you feel but not always: The empathic brain and its modulation. *Curr Opin Neurobiol* 18:153–158.
15. Anstey ML, Rogers SM, Ott SR, Burrows M, Simpson SJ (2009) Serotonin mediates behavioral gregarization underlying swarm formation in desert locusts. *Science* 323: 627–630.
16. Crockett MJ, Clark L, Tabibnia G, Lieberman MD, Robbins TW (2008) Serotonin modulates behavioral reactions to unfairness. *Science* 320:1739.
17. Way BM, Lačan G, Fairbanks LA, Melega WP (2007) Architectonic distribution of the serotonin transporter within the orbitofrontal cortex of the vervet monkey. *Neuroscience* 148:937–948.
18. Hornung J (2010) Handbook of the Behavioral Neurobiology of Serotonin. Elsevier, eds Muller CP, Jacobs BL (Elsevier, London), pp 51–64.
19. Koenigs M, Tranel D (2007) Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the Ultimatum Game. *J Neurosci* 27:951–956.
20. Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003) The neural basis of economic decision-making in the Ultimatum Game. *Science* 300:1755–1758.
21. Moll J, Schulkin J (2009) Social attachment and aversion in human moral cognition. *Neurosci Biobehav Rev* 33:456–465.
22. Crockett MJ (2009) The neurochemistry of fairness: Clarifying the link between serotonin and prosocial behavior. *Ann N Y Acad Sci* 1167:76–86.
23. Miczek KA, et al. (2007) Neurobiology of escalated aggression and violence. *J Neurosci* 27:11803–11806.
24. Davidson RJ, Putnam KM, Larson CL (2000) Dysfunction in the neural circuitry of emotion regulation—a possible prelude to violence. *Science* 289:591–594.
25. Krakowski M (2003) Violence and serotonin: Influence of impulse control, affect regulation, and social functioning. *J Neuropsychiatry Clin Neurosci* 15:294–305.
26. Ochsner KN, Gross JJ (2005) The cognitive control of emotion. *Trends Cogn Sci* 9:242–249.
27. Bechara A, Damasio H, Damasio AR (2000) Emotion, decision making and the orbitofrontal cortex. *Cereb Cortex* 10:295–307.
28. Damasio AR, Tranel D, Damasio H (1990) Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behav Brain Res* 41: 81–94.
29. Crockett MJ, Clark L, Robbins TW (2009) Reconciling the role of serotonin in behavioral inhibition and aversion: Acute tryptophan depletion abolishes punishment-induced inhibition in humans. *J Neurosci* 29:11993–11999.
30. Dayan P, Huys QJ (2009) Serotonin in affective control. *Annu Rev Neurosci* 32:95–126.
31. Cooper JR, Bloom FE, Roth RH, eds (2003) *The Biochemical Basis of Neuropharmacology* (Oxford Univ Press, New York), 8th Ed.
32. Chamberlain SR, et al. (2009) Atomoxetine modulates right inferior frontal activation during inhibitory control: A pharmacological functional magnetic resonance imaging study. *Biol Psychiatry* 65:550–555.
33. Robbins TW, Arnsten AF (2009) The neuropsychopharmacology of fronto-executive function: Monoaminergic modulation. *Annu Rev Neurosci* 32:267–287.
34. Greene JD (2007) Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends Cogn Sci* 11:322–323, author reply 323–324.
35. Huebner B, Dwyer S, Hauser M (2009) The role of emotion in moral psychology. *Trends Cogn Sci* 13:1–6.
36. Greene JD, Nystrom LE, Engell AD, Darley JM, Cohen JD (2004) The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44:389–400.
37. Huebner B, Hauser MD, Pettit P How the source, inevitability, and means of bringing about harm interact in folk-moral judgments. *Mind Lang*, in press.
38. Guth W, Schmittberger R, Schwarze B (1982) An experimental analysis of ultimatum bargaining. *J Econ Behav Organ* 3:367–388.
39. Camerer C (2003) *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton Univ Press, Princeton, NJ).
40. Pillutla MM, Murnighan JK (1996) Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organ Behav Hum Decis Process* 68:208–224.
41. Tabibnia G, Satpute AB, Lieberman MD (2008) The sunny side of fairness: Preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychol Sci* 19:339–347.
42. van't Wout M, Kahn RS, Sanfey AG, Aleman A (2006) Affective state and decision-making in the Ultimatum Game. *Exp Brain Res* 169:564–568.
43. Chapman HA, Kim DA, Susskind JM, Anderson AK (2009) In bad taste: evidence for the oral origins of moral disgust. *Science* 323:1222–1226.
44. Knoch D, Pascual-Leone A, Meyer K, Treyer V, Fehr E (2006) Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314:829–832.
45. Raleigh MJ, Brammer GL, McGuire MT, Yuwiler A (1985) Dominant social status facilitates the behavioral effects of serotonergic agonists. *Brain Res* 348:274–282.
46. Cools R, et al. (2005) Tryptophan depletion disrupts the motivational guidance of goal-directed behavior as a function of trait impulsivity. *Neuropsychopharmacology* 30:1362–1373.
47. Berman ME, McCloskey MS, Fanning JR, Schumacher JA, Coccaro EF (2009) Serotonin augmentation reduces response to attack in aggressive individuals. *Psychol Sci* 20: 714–720.
48. Davis MH (1980) A multidimensional approach to individual differences in empathy. *JSAS Catalog Sel Doc Psych* 10:85.
49. Singer T, et al. (2004) Empathy for pain involves the affective but not sensory components of pain. *Science* 303:1157–1162.
50. Knoch D, Fehr E (2007) Resisting the power of temptations: The right prefrontal cortex and self-control. *Ann N Y Acad Sci* 1104:123–134.
51. Ciaramelli E, Muccioli M, Ladavas E, di Pellegrino G (2007) Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Soc Cogn Affect Neurosci* 2:84–92.
52. Moretti L, Dragone D, di Pellegrino G (2009) Reward and social valuation deficits following ventromedial prefrontal damage. *J Cogn Neurosci* 21:128–140.
53. Krajchich I, Adolphs R, Tranel D, Denburg NL, Camerer CF (2009) Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *J Neurosci* 29:2188–2192.
54. Zahn R, et al. (2009) The neural basis of human social values: Evidence from functional MRI. *Cereb Cortex* 19:276–283.
55. Harmer CJ, et al. (2003) Acute SSRI administration affects the processing of social cues in healthy volunteers. *Neuropsychopharmacology* 28:148–152.
56. Besel LDS, Yuille JC (2010) Individual differences in empathy: The role of facial expression recognition. *Pers Individ Dif* 49:107–112.
57. Jorgensen H, Riis M, Knigge U, Kjaer A, Warberg J (2003) Serotonin receptors involved in vasopressin and oxytocin secretion. *J Neuroendocrinol* 15:242–249.
58. Insel TR (2010) The challenge of translation in social neuroscience: A review of oxytocin, vasopressin, and affiliative behavior. *Neuron* 65:768–779.
59. Deakin JF (2003) Depression and antisocial personality disorder: Two contrasting disorders of 5HT function. *J Neural Transm Suppl* 79–93.
60. Soderstrom H, Blennow K, Manhem A, Forsman A (2001) CSF studies in violent offenders. I. 5-HIAA as a negative and HVA as a positive predictor of psychopathy. *J Neural Transm* 108:869–878.
61. Frankle WG, et al. (2005) Brain serotonin transporter distribution in subjects with impulsive aggressivity: A positron emission study with [<sup>11</sup>C]McN 5652. *Am J Psychiatry* 162:915–923.
62. Coccaro EF, Lee RJ, Kavoussi RJ (2009) A double-blind, randomized, placebo-controlled trial of fluoxetine in patients with intermittent explosive disorder. *J Clin Psychiatry* 70:653–662.
63. Koenigs M, Kruepke M, Newman JP (2010) Economic decision-making in psychopathy: A comparison with ventromedial prefrontal lesion patients. *Neuropsychologia* 48: 2198–2204.
64. Berrios GE, et al. (1992) Feelings of guilt in major depression. Conceptual and psychometric aspects. *Br J Psychiatry* 160:781–787.
65. Chamberlain SR, et al. (2006) Neurochemical modulation of response inhibition and probabilistic learning in humans. *Science* 311:861–863.
66. Watson D, Clark LA, Tellegen A (1988) Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 54:1063–1070.
67. Hanley JA, Negassa A, Edwardes MDB, Forrester JE (2003) Statistical analysis of correlated data using generalized estimating equations: an orientation. *Am J Epidemiol* 157:364–375.