

# Vasopressin increases human risky cooperative behavior

Claudia Brunnlieb<sup>a,b,c,d,1</sup>, Gideon Nave<sup>e,1</sup>, Colin F. Camerer<sup>e,f,2</sup>, Stephan Schosser<sup>c</sup>, Bodo Vogt<sup>c,d</sup>, Thomas F. Münte<sup>a</sup>, and Marcus Heldmann<sup>a,b</sup>

<sup>a</sup>Department of Neurology, University Lübeck, 23562 Lübeck, Germany; <sup>b</sup>Center for Behavioral Brain Sciences, 39120 Magdeburg, Germany; <sup>c</sup>Department of Empirical Economics, Otto-von-Guericke-University Magdeburg, 39106 Magdeburg, Germany; <sup>d</sup>Department of Social Medicine and Health Economics, Otto-von-Guericke-University Magdeburg, 39120 Magdeburg, Germany; <sup>e</sup>Computation and Neural Systems, California Institute of Technology, Pasadena, CA 91125; and <sup>f</sup>Department of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125

Edited by Michael S. Gazzaniga, University of California, Santa Barbara, CA, and approved January 5, 2016 (received for review September 24, 2015)

**The history of humankind is an epic of cooperation, which is ubiquitous across societies and increasing in scale. Much human cooperation occurs where it is risky to cooperate for mutual benefit because successful cooperation depends on a sufficient level of cooperation by others. Here we show that arginine vasopressin (AVP), a neuropeptide that mediates complex mammalian social behaviors such as pair bonding, social recognition and aggression causally increases humans' willingness to engage in risky, mutually beneficial cooperation. In two double-blind experiments, male participants received either AVP or placebo intranasally and made decisions with financial consequences in the "Stag hunt" cooperation game. AVP increases humans' willingness to cooperate. That increase is not due to an increase in the general willingness to bear risks or to altruistically help others. Using functional brain imaging, we show that, when subjects make the risky Stag choice, AVP down-regulates the BOLD signal in the left dorsolateral prefrontal cortex (dlPFC), a risk-integration region, and increases the left dlPFC functional connectivity with the ventral pallidum, an AVP receptor-rich region previously associated with AVP-mediated social reward processing in mammals. These findings show a previously unidentified causal role for AVP in social approach behavior in humans, as established by animal research.**

vasopressin | intranasal administration | cooperation | fMRI | neuroeconomics

No other species shows the level of cooperative behavior that humans do. From hunter-gatherer communal sharing to spice trading to international response to catastrophes, an ever-increasing scale of cooperation marks the history of humankind and accounts for its successes (1). How have humans evolved to cooperate? Cooperation is often studied using behavioral paradigms where it is costly for an individual to cooperate, such as in social dilemmas in which prosocial choices benefit another individual at a cost to oneself (2). However, naturally occurring cooperation is often individually beneficial provided enough other individuals are also cooperative (called "risky cooperation") (3, 4). A recent theory, the "interdependent hypothesis," postulates that humans have evolved a unique skill for mutually beneficial cooperation in the context of especially risky coordination problems (5). The interdependent hypothesis is consistent with the risk sensitivity observed in primate group hunting in the wild (6, 7) with recent comparative studies reporting that humans are more cooperative than other primates only when cooperation requires increased willingness to bear social risks (8). As human cooperative behavior is associated with self-reported measures of pleasure and satisfaction (9, 10) and neural activation of the reward system (11, 12), it seems that our species has developed a biological mechanism that allows us to overcome the riskiness of cooperation to coordinate mutually beneficial social actions (13).

The development of neuromodulatory systems in the brain is a candidate biological mechanism underlying the evolution of mammalian social behaviors (14). Specifically, the neuropeptide arginine vasopressin (AVP) acts both as a neurotransmitter and a

hormone with widespread targets including regions of the prefrontal cortex, the amygdala, and the hippocampus and interacts with dopaminergic, reward-processing circuits in the ventral pallidum (VP) and nucleus accumbens (14). AVP regulates male mammalian social behaviors such as affiliation, aggression, monogamy, and paternal behaviors (14). Scholars have suggested that AVP mediates male pair bonding in mammals by stimulating dopamine release in the VP, inducing reward associations with the partner (15). Rodent studies have also linked AVP with social recognition, an important capacity required for distinguishing between conspecifics (16).

Like cooperation, many AVP-mediated mammalian social behaviors are potentially rewarding but are also risky: the benefits of spending all mating resources on a single female partner and engaging in paternal investment (at the absence of assured paternity), for example, depend on uncertain actions of conspecifics (17). Based on the proposition that AVP enhances the reward associated with social interactions in mammals (18), we hypothesized that AVP encodes the intrinsic value associated with cooperation in humans, overcoming the riskiness of social interactions to facilitate cooperation among humans.

AVP crosses the blood-brain barrier after intranasal administration, providing means for directly manipulating the level of AVP in the human central nervous system (19). Here we delineate the impact of intranasal AVP on risky cooperative behavior in a double-blind, placebo-controlled design, using a Stag hunt game, an experimental economic paradigm that is commonly used to study mutually beneficial risky cooperation (20, 21) (Fig. 1). In

## Significance

**Most forms of cooperative behavior take place in a mutually beneficial context where cooperation is risky as its success depends on unknown actions of others. In two pharmacological experiments, we show that intranasal administration of arginine vasopressin (AVP), a hormone that regulates mammalian social behaviors such as monogamy and aggression, increases humans' tendency to engage in mutually beneficial cooperation. Several control tasks ruled out that AVP's effects were driven by increased willingness to bare risks in the absence of social context, beliefs about the actions of one's partner, or altruistic concerns. Our findings provide novel causal evidence for a biological factor underlying cooperation and are in accord with previous findings that cooperation is intrinsically rewarding for humans.**

Author contributions: C.B., G.N., C.F.C., S.S., B.V., T.F.M., and M.H. designed research; C.B. performed research; C.B., G.N., C.F.C., B.V., T.F.M., and M.H. contributed new reagents/analytic tools; C.B., G.N., S.S., T.F.M., and M.H. analyzed data; and C.B., G.N., C.F.C., B.V., T.F.M., and M.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>C.B. and G.N. contributed equally to this work.

<sup>2</sup>To whom correspondence should be addressed. Email: camerer@hss.caltech.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1518825113/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1518825113/-DCSupplemental).

the Stag hunt, two players simultaneously choose between a cooperative action that potentially yields a high payoff (A, “Stag”) and a noncooperative action (B, “Rabbit”), the payoffs of which are lower but less risky (the choice labels “Stag” and “Rabbit” originate from a stylized description of the choice between safe individual hunting for low-value Rabbits, compared with hunting for a high-value Stag that requires many individuals to collaborate in the hunt). By choosing Stag, a player may increase his or her payoff, but also risks ending up worse off if his partner did not cooperate by choosing Stag.

The Stag hunt game has two stable behavioral patterns that are mutually optimal for both players (called “pure Nash equilibria” in game theory). If one’s partner is sufficiently likely to choose Stag, then Stag choice is optimal for an individual. The (Stag, Stag) choice pair is therefore a self-fulfilling equilibrium pattern and earns the highest payoffs for both players (Fig. 1A, top left of the payoff matrix). In the other equilibrium pattern, both players fear that the partner will not cooperate and hence choose the safer Rabbit strategy (Fig. 1A, bottom right of the payoff matrix). Note that (Stag, Stag) groups earn more, collectively, than (Rabbit, Rabbit) groups. Thus, the rate of Stag choice indicates whether a group has a self-fulfilling norm of beneficial social risk-taking. Crucially, the Stag hunt game is different from a social dilemma game (e.g., the prisoner’s dilemma or the public goods game) because even in a single play of the game, the cooperative Stag choice is individually optimal if one’s partner is likely to choose Stag (on the contrary, defection is the optimal strategy of a selfish player facing a dilemma, regardless of the partner’s behavior).

We manipulated the social risk associated with a Stag hunt game by varying its *security level* (denoted “*S*”). The security level *S* is the minimal payoff a player could secure himself by choosing Rabbit (Fig. 1B). A high security level makes the Rabbit action more attractive; thus, Stag becomes less attractive and more risky, as the likelihood that one’s partner will choose Stag presumably decreases. We hypothesized that AVP will increase the rates of Stag choice in the Stag hunt game and also tested whether any effect of AVP is due to other measurable factors that could potentially influence Stag choices, namely, subjects’ general attitude toward financial risk, the belief about the partners’ strategy choice, and social preferences (i.e., willingness to incur a cost to benefit others). Finally, we investigate the neural mechanisms underlying AVP’s effects on human cooperative behavior using functional magnetic resonance imaging (fMRI).

### Experiment 1 (Behavior)

Fifty-nine healthy adult males received nasal spray with either 40 International Units (IU) of AVP or placebo, 30 min before the task (19). We presented subjects with a Stag hunt game matrix (Fig. 1B) and asked them to choose (22) the minimal security level at which they would switch from the safe Rabbit action to the risky Stag action (we used the abstract labels “A” and “B”

rather than Stag and Rabbit to avoid possible framing effects). Then, each subject was randomly matched with a partner, and a value of *S* (between 0 and 150) was realized at random. Based on the drawn *S* value and the subjects’ switching points, we determined whether the subjects chose the Stag or Rabbit action and calculated their corresponding payoffs.

The hypothesis that AVP increases cooperative behavior implies that subjects under intranasal AVP would be willing to cooperate even when the security level associated with the Rabbit choice were greater—that is, they would choose a higher switching point than the placebo group. In line with this hypothesis, the AVP group switched from Rabbit to Stag at a higher security level on average [Fig. 2A; one-way ANOVA,  $F(1,57) = 5.522, P < 0.023$ ].

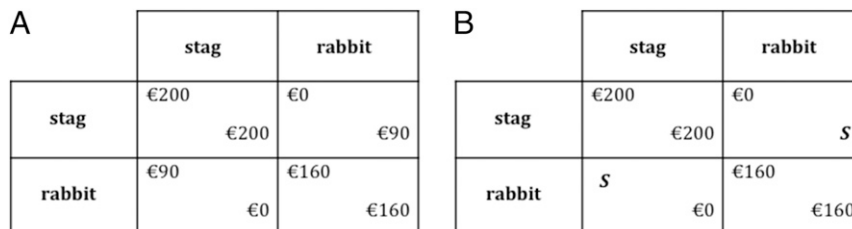
Several factors might underlie the increased cooperation rates following AVP treatment. First, because Stag payoffs have higher variance (i.e., they are riskier), AVP might simply increase the willingness to accept risky financial payoffs, even absent social context. To test this alternative hypothesis, the Stag hunt game was followed by a task in which the same subjects (still under AVP influence) chose between lotteries with different financial risk levels and no social interdependency (*SI Appendix*, 1.1.3). In this task, where payoffs were kept constant relative to the Stag hunt game but the risk associated with the different options was not socially generated (i.e., was not associated with uncertainty regarding the partner’s actions), the AVP group was *not* more likely to take riskier actions than the placebo group [*SI Appendix*, Table S5; one-way ANOVA,  $F(1,56) = 0.339, P = 0.563$ ].

Second, AVP-induced cooperation could be mediated by changes in social preferences (i.e., caring about the welfare of one’s partner) (23, 24). To measure social preference, the same subjects took part in another task where they chose between five possible monetary allocations that embodied systematic trade-off between the payoffs to oneself and to another person (*SI Appendix*, Table S3). In line with a previous study (25), there was no effect of AVP on social preference [*SI Appendix*, Fig. S2; one-way ANOVA,  $F(1,57) = 1.242, P = 0.270$ ].

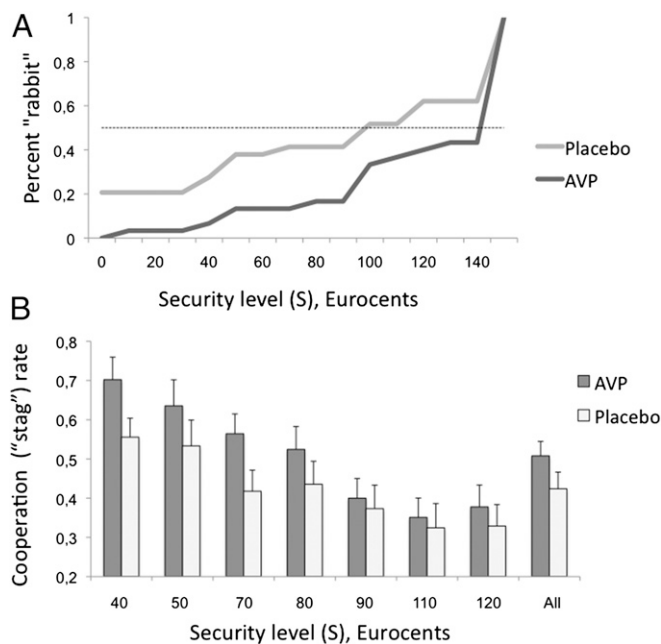
Third, AVP could also influence one’s beliefs regarding his partner’s actions. We elicited the subjects’ beliefs about the likelihood that their partner chose the cooperative strategy by presenting them 12 Stag hunt games in a pseudorandomized order and asking them to indicate the likelihood (between 0 and 100%) that their partner chose “Stag.” Again, we found no reliable AVP effect (*SI Appendix*, Fig. S1 and Table S6).

Fourth, we ruled out the possibility that AVP increased cooperation is due to altered self-reported mood [pleasantness:  $F(1,57) = 0.48, P = 0.490$ ; wakefulness:  $F(1,57) = 1.27, P = 0.264$ ; calmness:  $F(1,57) = 0.24, P = 0.629$ ] (*SI Appendix*, 1.2.5).

Our findings demonstrate that AVP’s effects on cooperation are specific to the mutually beneficial social context of the Stag hunt game and are consistent with the notion that AVP increases



**Fig. 1.** The Stag hunt task. (A) Stag hunt game payoff matrix. The cooperative action Stag potentially generates the highest possible payoff, but only if one’s partner cooperates as well. Rabbit choice secures a lower (but certain) payoff, regardless of the partner’s action. The payoff dominant equilibrium (200/200) is at the *Top Left*, and the risk-dominant equilibrium (160/160) is at the *Bottom Right*. (B) Stag hunt game payoff matrix of the strategy method. Participants indicated a security level (parameter *S*-value) at which they would switch between the risk-dominant Rabbit strategy to the payoff-dominant Stag strategy and also indicated the direction of switching.



**Fig. 2.** AVP increased cooperative behavior. (A) AVP subjects switched from Stag to Rabbit at a higher security level (*S*-value), indicating higher cooperation rates and reduced aversion to social risk under AVP treatment. Note that, as all subjects switched from Stag to Rabbit, the cumulative distribution was equivalent to the rate of Rabbit choice at any given security level. (B) Mean cooperation rates of both groups in experiment 2 (fMRI). AVP subjects were significantly more cooperative; the effect increased when the incentive to cooperate was high.

the expected reward of cooperation, rather than general risk attitudes, social preferences, or beliefs about the partner's actions.

### Experiment 2 (Behavior and Brain Imaging)

To further investigate the neural mechanisms underlying AVP's effect on risky cooperation, we conducted a second Stag hunt experiment using fMRI. We administered AVP or placebo to 34 healthy adult males in a randomized double-blind regime. Before scanning, each subject was introduced to a fixed partner who, unbeknownst to the subject, was a confederate of the experimenter. Participants played 105 different games derived from seven versions of the Stag hunt game that differed only in their security level (*Materials and Methods* and *SI Appendix, Tables S8 and S9*). Subjects made their choices by pressing a button and received no feedback about their partner's actions and payoffs, to disable the complicated effects of learning and reputation building.

A mixed-effect logistic model with subjects' decisions as the dependent variable (*SI Appendix, Table S10*) revealed significant effects of security level ( $z = -6.925$ ,  $P < 0.001$ ), drug ( $z = 2.571$ ,  $P < 0.01$ ), and a drug  $\times$  security level interaction ( $z = -2.103$ ,  $P < 0.03$ ). These results replicate the behavioral results of study 1 (Fig. 2B). Furthermore, a mixed-effect model with players' response time (RT) as the dependent variable resulted in a significant drug  $\times$  choice interaction. The AVP group had a significantly faster RT when choosing Stag compared with Rabbit ( $t = -2.144$ ,  $P < 0.03$ ) (Fig. 3A and *SI Appendix, Table S11*).

We estimated cooperation-related brain activity using a general linear model (GLM) on functional MR images measuring blood oxygen level-dependent activity (BOLD signal). The GLM included dummy regressors for actual Stag and Rabbit choices and nuisance parameters controlling for head motion (*Materials and Methods*). We calculated the contrast between BOLD activity from

Stag vs. Rabbit choices at the first (subject) level and entered resulting *t*-contrasts of all subjects into two-sample *t* tests by comparing AVP group vs. placebo group and vice versa.

This between-group comparison revealed increased activation of the left dorsolateral prefrontal cortex (dlPFC) (Fig. 3B) in the placebo group relative to the AVP group when making Stag vs. Rabbit choices. An additional region-of-interest analysis of the cluster showed that the AVP group had left dlPFC activity that was reduced during Stag choices ( $t = -2.34$ ,  $P < 0.05$ ) and heightened during Rabbit choices ( $t = 1.87$ ,  $P < 0.05$ ) (Fig. 3C), relative to the placebo group. Using  $\beta$ -series correlation based functional connectivity analyses (*Materials and Methods*), we found altered functional connectivity of the left dlPFC with the ventrolateral PFC (vlPFC) and the parahippocampal gyrus, depending on the strategy by drug treatment interaction (*SI Appendix, Table S15* and Figs. S4 and S5). We found that under AVP influence the left dlPFC showed increased functional connectivity with the left VP, the cingulate gyrus, and the medial and superior frontal gyrus during Stag choices and with the left parahippocampal gyrus, left amygdala, and anterior cingulate cortex (ACC) during Rabbit choices (Fig. 3D and *SI Appendix, Tables S13* and S14).

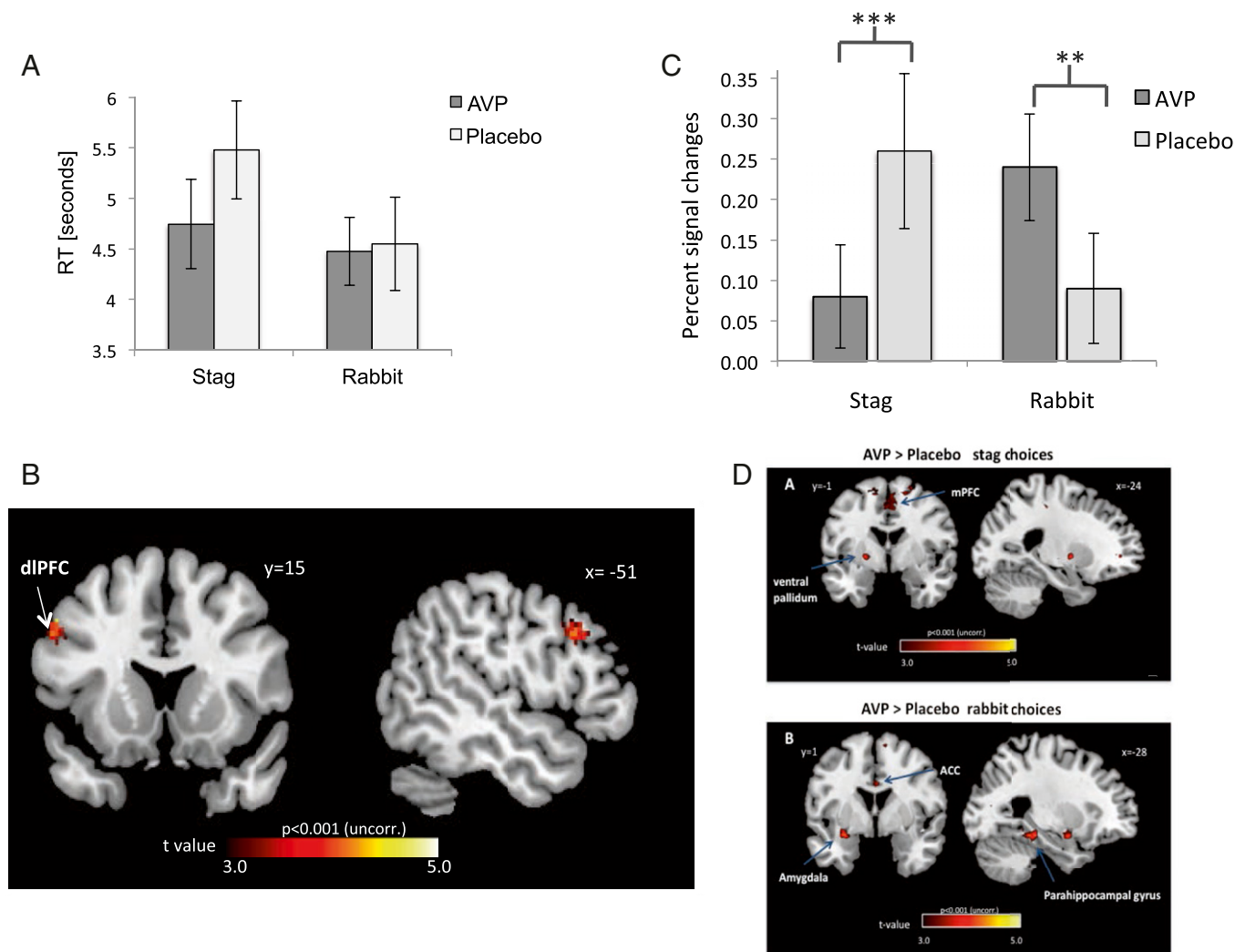
### Discussion

The current study investigated the effect of intranasally administered AVP on mutually beneficial cooperative behavior in humans. We provide, to our knowledge, the first evidence that AVP increases risky cooperative behavior and that this effect is specific to the mutually beneficial context of the Stag hunt game and not mediated by general risk attitudes, a change in beliefs, social preferences, or mood.

Although we acknowledge that inferring cognitive processes from fMRI activation patterns is not straightforward (26), our fMRI investigation suggests a possible neural mechanism underlying AVP-induced cooperation. Given the established links between increased left dlPFC activation and cognitive control (27–29) and the known role of the left dlPFC in risk integration (meta analysis) (30), high-level planning, and behavioral inhibition (31), we speculate that AVP administration has reduced the need to recruit this cortical region to inhibit the choice of the noncooperative Rabbit action during cooperation.

Indeed, Rabbit seems to be the default appetitive choice in the placebo group (as evident by the choice frequencies), but under AVP treatment cooperation (Stag) becomes the default strategy. As a consequence, choosing it requires less processing time (as indicated by the reaction times) and less cognitive control (as indicated by reduced left dlPFC activity) following AVP treatment. Moreover, because under AVP treatment Rabbit is no longer the default strategy, choosing it requires additional cognitive control from dlPFC, coupled with brain activity in regions that are associated with inhibition of prepotent suboptimal responses (vlPFC) (31), conflict (ACC) (32), and emotional vigilance (amygdala and parahippocampus gyrus) (33).

The enhanced functional coupling of the left dlPFC and VP during cooperative choices in the AVP group provides further support for the hypothesis that AVP increases the intrinsic expected reward associated with cooperation. The VP is part of the basal ganglia reward circuitry and has a high density of both AVP V1a and dopamine receptors (14). Recently, scholars have suggested that the interaction of the dopaminergic and vasopressinergic systems in the VP encodes the rewarding properties of social interactions, facilitating social recognition and pair bonding in several mammalian species (15, 34). Our data suggest that the vasopressinergic/dopaminergic interactions in the reward system (14) are evolutionarily stable across species and might be linked to a broader range of social behaviors as established by animal research so far.



**Fig. 3.** AVP caused choice-dependent changes in brain activity. (A) Mean response times across action choices. The placebo group was significantly slower when choosing Stag compared with Rabbit (paired-sample  $t = 2.62$ ,  $P < 0.01$ ); the effect disappeared in the AVP group ( $t = -2.144$ ,  $P < 0.03$ ). (B) A cluster in the left dIPFC showed increased BOLD activation for the contrast Stag > Rabbit choices in the placebo group relative to the AVP group. (C) A region-of-interest analysis in the area shown in B revealed a significant drug  $\times$  choice interaction. AVP decreased the BOLD signal (percentage signal change, y axis) in the left dIPFC during Stag choices ( $t = -2.34$ ,  $P < 0.05$ ) and increased activity in the latter brain region during Rabbit choices ( $t = 1.87$ ,  $P < 0.05$ ), relative to the placebo group. This finding indicates that cooperation (choosing Stag) required less mental effort under AVP treatment, in accordance with the behavioral and RT data. (D) Intranasal AVP enhanced the left dIPFC functional connectivity with (i) the left ventral pallidum, the cingulate gyrus, and the medial and superior frontal gyrus during Stag choices and (ii) increased the left dIPFC functional connectivity with the left parahippocampal gyrus, left amygdala, and ACC during Rabbit choices.

Furthermore, it has been argued that cooperation is intrinsically rewarding for humans (11–13) and that our unique capability to cooperate has evolved under environmental challenges selecting for adaptations to solve risky coordination problems (5, 8). These hypotheses are in accordance with our finding that AVP facilitates mutually beneficial risky cooperation by reducing aversion to social risk, without changing general aversion to risky financial outcomes or social preference.

Our experiments delineated AVP's role in mutually beneficial cooperation in an environment where altruistic concerns, reputation building, and strategically learning the partner's intentions and beliefs did not play a role (35, 36). Most importantly, AVP did not change behavior in the social preferences control task, which rules out AVP working purely through increased altruistic one-sided prosociality. However, previous work has shown that AVP administration increased cooperation following a cooperative gesture by one's partner in a repeated prisoner's dilemma (37), where cooperation might be mutually beneficial

(38)—in line with the interdependent hypothesis (5). We believe that further investigation of AVP's role in social behavior would benefit from using computational theories that distinguish the different mechanism through which AVP may foster cooperative behavior in repeated interactions (35), particularly belief inference (33, 36, 39) and inferential limits in disorder (40).

In summary, we combined insights and methodologies from biology, anthropology, and economics and documented a biological cause increasing cooperation—a behavioral feature that largely distinguishes humans from other species and is ubiquitous across human societies. Our fMRI results hint for an AVP interaction with cortical control and dopaminergic reward pathways in the VP. Further investigation of this interaction may have clinical implications for treatment of neuropsychiatric disorders associated with social malfunction. Social anxiety and attention deficit disorders, for example, are characterized by low dopamine levels (41–44), and animal research has reported that AVP stimulates dopamine release in the reward system (15, 34). This suggests that intranasally

administered AVP may also have the potential to promote dopamine release in human patients suffering from social disorders and thus may have substantial public health significance.

Finally, we highlight that human AVP research is in its early days. We hope that future investigations will further examine whether the effects are generalized to females (45) and other subject populations and allow translating decades of fruitful animal research into biologically informed theories of social behavior in humans.

## Materials and Methods

### Experiment 1 (Behavior).

**Subjects.** Fifty-nine healthy adult males [age = 19–32, mean = 24.3, SD = 2.9; AVP: mean = 24.4, SD = 2.9; placebo: mean = 24.1, SD = 2.9/ $t(57) = -0.501$ ,  $P = 0.618$ ] participated in this study. Exclusion criteria were psychiatric or neurological disorders, kidney disease, cardiovascular problems, asthma, or migraine. All subjects were right-handed. All sessions took place between 8:00 AM and 6:00 PM. Experiments were approved by the ethical committee of the University of Magdeburg and conducted in accordance with the Declaration of Helsinki.

**Experimental procedures and materials.** Participants were randomly allocated to isolated semicubicles, where they could not see or interact with each other. Each participant read and signed an informed consent. Participants self-administered a nasal spray with 40 IU of AVP or a placebo (saline) under the supervision of the experimenter 30 min prior to the experimental procedure. The time frame was chosen according to a previous study (19) that investigated the pharmacokinetics of intranasal AVP in the cerebrospinal fluid and indicated a high level within 15 min that kept increasing up to 80 min following the procedure. Participants completed a series of unrelated tests beginning with the filling of the German version of the Multidimensional Mood state questionnaire (MDMQ) (46). Afterward participants played the Stag hunt game in the strategy method (Fig. 1B and *SI Appendix*, 1.1.1) where they were asked to indicate a value of the parameter  $S$  in which they would switch from the Rabbit to the Stag strategy.

Next, we elicited participants' beliefs regarding their partners' actions by presenting them 12 Stag hunt games with different security levels (*SI Appendix*, Table S1). We incentivized the participants to indicate the likelihood (between 0 and 100%) that their partner chose "Stag" (see *SI Appendix*, 1.1.2 for further details).

Participants also took part in a lottery choice task that was designed to be identical to the Stag hunt task, except that the player's risk was not generated through a social interaction: subjects had to choose between a risky lottery A and a less risky lottery B of nine different lottery pairs. The earlier participants switched from the less risky lottery B to the risky lottery A, the more risk-seeking they were; later switching points thus indicate a more risk-averse behavior.

We measured social (distributional) preferences by asking all participants to choose one of five different monetary allocations that posited a trade-off between the payoff to one's self and one's anonymous partner. At the end of the experiment participants again filled out the MDMQ.

**Data analysis.** We tested for between-group differences in participants' Stag hunt game-switching decisions ( $S$ -values) using a one-way ANOVA. For the beliefs regarding the partner's actions in the 12 games, we estimated a mixed model with security level and treatment as fixed effects, clustered at the subject level. For the lottery task, we determined the switching point between the risky and the safe lotteries for each subject and entered them into a one-way ANOVA. We used the same test for the analysis of social preferences.

MDMQ scores were analyzed using mixed-effect general linear models with pre- and postexposure as a within-subject factor and the experimental condition (AVP vs. placebo) as a between-subject factor.

### Experiment 2 (Behavior and Brain Imaging).

**Subjects.** Thirty-four healthy adult male participants [age = 19–34, mean = 25.6, SD = 4.2, AVP: mean = 25.7, SD = 2.9/placebo: mean = 25.5, SD = 5.3/ $t(32) = 0.121$ ,  $P = 0.905$ ] participated in the study. Three participants were excluded from further analysis due to extensive head movements during scanning. One participant felt uncomfortable during the scanning procedure and did not complete the task. Thus, 30 participants were included in data analyses (15 from each treatment group).

**Experimental procedures and materials.** After signing informed consent agreements, participants self-administered (under the supervision of the experimenter) either a nasal spray with 20 IU of AVP or a placebo (saline) under a double-blind protocol. Because the total scanning procedure with premeasures (e.g., anatomical scan) and experimental procedure lasted ~50 min, intranasal administration was conducted 15 min prior to the experimental procedure to

ensure that AVP levels were elevated throughout the experiment, in line with the pharmacokinetic properties of intranasal AVP administration (19).

Before scanning, each participant was introduced to a fixed partner (the same person in all of the experimental sessions) who, unbeknownst to him, was a confederate of the experimenter. Participants were told that the partner would sit in another room and that both players would interact via a computer network. Participants were presented the Stag-hunt game matrix (Fig. 1), were instructed to act as row players, and were told that their partner would act as the column player.

Before the task, participants were familiarized with the Stag hunt game using a tutorial that was conducted outside the scanner. Only when the participants had fully understood the task did the experimental procedure begin. Each game period started with a fixation cross (1–2 s), and then the Stag hunt game matrix was presented (14 s) (*SI Appendix*, Fig. S3); participants had to indicate their choice of strategy ("Stag," "Rabbit") by a button press (the names "Stag" and "Rabbit" were replaced by the abstract labels "A" and "B" in the task). In total, 105 different Stag hunt games were presented and the task lasted about 28 min. We used seven different variations of the Stag hunt game ("basis games") that were identical with respect to the values of all payoffs and differed only in their security level, parameter  $S$  (*SI Appendix*, Table S8). Overall, subjects played each of the basic games 15 times (a total of 105 games) in a pseudorandomized order. To avoid stimulus adaptation, each of the seven basic Stag hunt games was varied in its payoffs 14 additional times by increasing the entire payoff matrix by integer multiples of 10 euro cents (*SI Appendix*, Table S9). After the experiment, participants were paid for randomly chosen 4 of the 105 Stag hunt games, realized by choosing four sheets of paper from a box containing the numbers 1–105. Participants' choices were then matched with the choices of the confederate for the four games to determine the payoff for the participant.

**Behavioral data analysis.** We estimated a logistic regression mixed model where the dependent variable was the choice of strategy (1 = Stag, 0 = Rabbit), and the independent variables were drug (1 = AVP, 0 = placebo), security level (parameter  $S$ ), and the interaction between them, clustered at the subject level. RTs were analyzed using a linear regression mixed model, where the independent variables were treatment (1 = AVP, 0 = placebo), choice (1 = Stag, 0 = Rabbit), security level, security level  $\times$  choice interaction, and drug  $\times$  choice interaction, clustered at the subject level.

**Standard fMRI analysis.** Analyses of fMRI data were conducted using the *SPM8* software package (Wellcome Department of Imaging Neuroscience, University College London). Preprocessing included slice time correction, motion correction, coregistration, spatial normalization, and spatial smoothing (using a Gaussian kernel with a full width at half maximum of 8 mm). Additionally, we applied a high-pass temporal filtering (128 s) and estimated a GLM that included separate boxcar regressors for the choices of Stag and Rabbit. These regressors had the value of +1 for the entire duration of the decision epoch. Movement parameters ( $x$ ,  $y$ ,  $z$ , pitch, roll, and yaw) were included in the model to minimize movement-correlated effects. To control for serial correlations, we applied the standard SPM AR (1) autoregressive model. At the first level GLM, all Stag choices were weighted +1, and Rabbit choices were weighted -1 to reveal brain regions in which activity was significantly greater during Stag vs. Rabbit choices. Resulting  $t$ -contrast maps of all subjects were entered into two-sample  $t$ -tests. These contrasts were considered at a threshold of  $P < 0.001$  (uncorrected) and a voxel extent threshold of  $k = 10$ . The contrast placebo > AVP for Stag > Rabbit choices revealed increased BOLD activation in the left dlPFC (*SI Appendix*, Table S12). We further conducted a functional region-of-interest (ROI) analysis with the functional cluster of the left dlPFC (peak coordinates:  $x = -54$ ,  $y = 18$ ,  $z = 34$ ) from the between-group comparison placebo > AVP (Stag > Rabbit choices). The ROI was created using *marsbar* toolbox and entered into the *rfx-plot* toolbox.

**Functional connectivity analyses.** We performed functional connectivity analyses using the method by Rissman (47) to analyze how AVP affects the interaction between the left dlPFC and other brain regions during Stag and Rabbit choices. We used the Rissman method because a recent investigation (48) has concluded that the method is more sensitive for the case of event-related designs with more trial repetitions (as the case in our experiment) and retains more power under conditions of hemodynamic response function variability.

For each participant, we estimated another GLM using separate covariates to model BOLD responses for Stag and Rabbit choices in every single trial and extracted  $\beta$ -values of both experimental conditions to calculate condition-specific  $\beta$ -series. According to this method, brain regions in which  $\beta$ -series coefficients are correlated under a given condition exhibit task-related functional connectivity (47). The functional cluster of the left dlPFC that showed significant activation for the contrast placebo > AVP for Stag > Rabbit choices was used as a seed region (peak coordinates:  $x = -54$ ,  $y = 18$ ,  $z = 34$ ). We estimated a  $\beta$ -series within the latter cluster, averaged across voxels, and calculated the correlation of the  $\beta$ -series with the  $\beta$ -series of every other voxel in the brain—resulting in separate

correlation maps for Stag and Rabbit for each participant (subject level). Finally, we normalized the correlation maps using an arc-hyperbolic tangent transformation (note that the normalized correlations are called “correlations” by convention, but they are not bounded by  $-1$  and  $+1$ ) and entered the normalized correlation maps into two-sample  $t$ -tests. We conducted the comparison  $AVP > placebo$  for Stag choices at a significance level of  $P < 0.001$  (uncorrected cluster level = 10; see correlation map in Fig. 3D and *SI Appendix, Table S13*), whereas for the reverse contrast ( $placebo > AVP$ ) we found no brain region at the chosen threshold. Regarding Rabbit choices, we conducted the comparison

$AVP > placebo$  at  $P < 0.001$  (uncorrected; cluster level = 10; see correlation map in Fig. 3D and *SI Appendix, Table S14*). Furthermore, we calculated a drug ( $AVP, placebo$ )  $\times$  condition (Stag, Rabbit) contrast using the flexible factorial design matrix of *SPM8* (*SI Appendix, Figs. S4 and S5 and Table S15*).

**ACKNOWLEDGMENTS.** This work was supported by a special grant of the Center for Behavioral Brain Sciences, Magdeburg (to M.H.), Deutsche Forschungsgemeinschaft (DFG) (T.M.), and the Betty and Gordon Moore Foundation (C.F.C. and G.N.).

- Boyd R, Richerson PJ (2009) Culture and the evolution of human cooperation. *Philos Trans R Soc Lond B Biol Sci* 364(1533):3281–3288.
- Cooper R, DeJong DV, Forsythe R, Ross TV (1996) Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games Econ Behav* 12(2):187–218.
- Smith, Adam (1991) *The Wealth of Nations*, ed Skinner AS (Prometheus Books, New York), Vol 3.
- Skyrms B (2004) *The Stag Hunt and the Evolution of Social Structure* (Cambridge University Press, Cambridge, UK).
- Tomasello M, Melis AP, Tennie C, Wyman E, Herrmann E (2012) Two key steps in the evolution of human cooperation the interdependence hypothesis. *Curr Anthropol* 53(6):673–692.
- Watts DP, Mitani JC (2002) Hunting behavior of chimpanzees at Ngogo, Kibale National Park, Uganda. *Int J Primatol* 23(1):1–28.
- Gilby IC, Wrangham RW (2007) Risk-prone hunting by chimpanzees (*Pan troglodytes schweinfurthii*) increases during periods of high diet quality. *Behav Ecol Sociobiol* 61(11):1771–1779.
- Duguid S, Wyman E, Bullinger AF, Herfurth-Majstorovic K, Tomasello M (2014) Coordination strategies of chimpanzees and human children in a Stag Hunt game. *Proc Biol Sci* 281(1796):20141973.
- Gächter S, Fehr E (2002) *Fairness in the Labour Market: A Survey of Experimental Results* (Physica-Verlag HD), pp 95–132.
- Haselhuber MP, Mellers BA (2005) Emotions and cooperation in economic games. *Brain Res Cogn Brain Res* 23(1):24–33.
- Decety J, Jackson PL, Sommerville JA, Chaminade T, Meltzoff AN (2004) The neural bases of cooperation and competition: An fMRI investigation. *Neuroimage* 23(2):744–751.
- Rilling J, et al. (2002) A neural basis for social cooperation. *Neuron* 35(2):395–405.
- Tabibnia G, Lieberman MD (2007) Fairness and cooperation are rewarding: Evidence from social cognitive neuroscience. *Ann N Y Acad Sci* 1118:90–101.
- Choleris D, Pfaff W, Kavaliers M, eds (2013) *Oxytocin, Vasopressin and Related Peptides in the Regulation of Behavior* (Cambridge University Press, Cambridge, UK).
- Lim MM, Young LJ (2004) Vasopressin-dependent neural circuits underlying pair bond formation in the monogamous prairie vole. *Neuroscience* 125(1):35–45.
- Bluthe RM, Schoenen J, Dantzer R (1990) Androgen-dependent vasopressinergic neurons are involved in social recognition in rats. *Brain Res* 519(1-2):150–157.
- Reichard UH, Boesch C (2003) *Monogamy: Mating Strategies and Partnerships in Birds, Humans and Other Mammals* (Cambridge University Press, Cambridge, UK).
- Freeman SM, Young L (2013) Oxytocin, vasopressin, and the evolution of mating systems in mammals. *Oxytocin, Vasopressin and Related Peptides in the Regulation of Behavior*, eds. E. Choleris, D. W. Pfaff, and M. W. Kavaliers (Cambridge University Press, Cambridge, UK), pp. 128–144.
- Born J, et al. (2002) Sniffing neuropeptides: A transnasal approach to the human brain. *Nat Neurosci* 5(6):514–516.
- Brosnan SF, Wilson BJ, Beran MJ (2012) Old World monkeys are more similar to humans than New World monkeys when playing a coordination game. *Proc Biol Sci* 279(1733):1522–1530.
- Camerer C (2003) *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton University Press, Princeton, NJ).
- Brandts J, Charness G (2011) The strategy versus the direct-response method: A first survey of experimental comparisons. *Exp Econ* 14:375–398.
- Charness G, Rabin M (2002) Understanding social preferences with simple tests. *Q J Econ* 117(3):817–869.
- Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Q J Econ* 114(3):817–868.
- Israel S, Weisel O, Ebstein RP, Bornstein G (2012) Oxytocin, but not vasopressin, increases both parochial and universal altruism. *Psychoneuroendocrinology* 37(8):1341–1344.
- Poldrack RA (2006) Can cognitive processes be inferred from neuroimaging data? *Trends Cogn Sci* 10(2):59–63.
- Braver TS, Cohen JD (2001) Working memory, cognitive control, and the prefrontal cortex: Computational and empirical studies. *Cogn Process* 2:25–55.
- Carter CS, van Veen V (2007) Anterior cingulate cortex and conflict detection: An update of theory and data. *Cogn Affect Behav Neurosci* 7(4):367–379.
- Hare TA, Camerer CF, Rangel A (2009) Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324(5927):646–648.
- Mohr PN, Biele G, Heekeren HR (2010) Neural processing of risk. *J Neurosci* 30(19):6613–6619.
- Mitchell DG (2011) The nexus between decision making and emotion regulation: A review of convergent neurocognitive substrates. *Behav Brain Res* 217(1):215–231.
- Botvinick MM, Cohen JD, Carter CS (2004) Conflict monitoring and anterior cingulate cortex: An update. *Trends Cogn Sci* 8(12):539–546.
- Bhatt MA, Lohrenz T, Camerer CF, Montague PR (2012) Distinct contributions of the amygdala and parahippocampal gyrus to suspicion in a repeated bargaining game. *Proc Natl Acad Sci USA* 109(22):8728–8733.
- Nair HP, Young LJ (2006) Vasopressin and pair-bond formation: Genes to brain to behavior. *Physiology (Bethesda)* 21:146–152.
- Camerer CF, Ho TH, Chong JK (2002) Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *J Econ Theory* 104(1):137–188.
- Yoshida W, Dolan RJ, Friston KJ (2008) Game theory of mind. *PLOS Comput Biol* 4(12):e1000254.
- Rilling JK, et al. (2012) Effects of intranasal oxytocin and vasopressin on cooperative behavior and associated brain activity in men. *Psychoneuroendocrinology* 37(4):447–461.
- Ely JC, Välimäki J (2002) A robust folk theorem for the prisoner's dilemma. *J Econ Theory* 102(1):84–105.
- Yoshida W, Seymour B, Friston KJ, Dolan RJ (2010a) Neural mechanisms of belief inference during cooperative games. *J Neurosci* 30(32):10744–10751.
- Yoshida W, et al. (2010b) Cooperation and heterogeneity of the autistic mind. *J Neurosci* 30(26):8815–8818.
- Iversen SD, Iversen LL (2007) Dopamine: 50 years in perspective. *Trends Neurosci* 30(5):188–193.
- Swanson JM, et al. (2007) Etiologic subtypes of attention-deficit/hyperactivity disorder: Brain imaging, molecular genetic and environmental factors and the dopamine hypothesis. *Neuropsychol Rev* 17(1):39–59.
- Tiihonen J, et al. (1997) Dopamine reuptake site densities in patients with social phobia. *Am J Psychiatry* 154(2):239–242.
- Schneier FR, et al. (2000) Low dopamine D(2) receptor binding potential in social phobia. *Am J Psychiatry* 157(3):457–459.
- Feng C, et al. (2015) Oxytocin and vasopressin effects on the neural response to social cooperation are modulated by sex in humans. *Brain Imaging Behav* 9(4):754–764.
- Steyer R, Schwenkmezger P, Notz P, Eid M (1997) *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF)*. *Handanweisung* (Hogrefe, Göttingen).
- Rissman J, Gazzaley A, D'Esposito M (2004) Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage* 23(2):752–763.
- Cisler JM, Bush K, Steele JS (2014) A comparison of statistical methods for detecting context-modulated functional connectivity in fMRI. *Neuroimage* 84(84):1042–1052.